

VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

Lina PAPŠIENĖ

ERDVINĖS INFORMACIJOS MODELIAVIMO
IR HARMONIZAVIMO TOBULINIMAS
SKAITMENINĖJE KARTOGRAFIJOJE

DAKTARO DISERTACIJA

TECHNOLOGIJOS MOKSLAI,
MATAVIMŲ INŽINERIJA (10T)



Vilnius LEIDYKLA
TECHNIKA 2014

Disertacija rengta 2010–2014 metais Vilniaus Gedimino technikos universitete.

Vadovai:

prof. dr. Eimuntas Kazimieras PARŠELIŪNAS (Vilniaus Gedimino technikos universitetas, matavimų inžinerija – 10T), (2012–2014),
prof. habil. dr. Vytautas GINIOTIS (Vilniaus Gedimino technikos universitetas, matavimų inžinerija – 10T), (2010–2012).

Vilniaus Gedimino technikos universiteto Matavimų inžinerijos mokslo krypties disertacijos gynimo taryba:

Pirmininkas

prof. habil. dr. Vladas VEKTERIS (Vilniaus Gedimino technikos universitetas, matavimų inžinerija – 10T).

Nariai:

doc. dr. Domantas BRUČAS (Vilniaus Gedimino technikos universitetas, matavimų inžinerija – 10T),
prof. dr. Marija BURINSKIENĖ (Vilniaus Gedimino technikos universitetas, statybos inžinerija – 02T),
doc. dr. Egidijus RIMKUS (Vilniaus universitetas, fizinė geografija – 06P),
habil. dr. Pawel Aleksander WIELGOSZ (Varmijos ir Mozūrijos universitetas Olštyne, matavimų inžinerija – 10T).

Disertacija bus ginama viešame Matavimų inžinerijos mokslo krypties disertacijos gynimo tarybos posėdyje **2014 m. gruodžio 18 d. 14 val.** Vilniaus Gedimino technikos universiteto senato posėdžių salėje.

Adresas: Saulėtekio al. 11, LT-10223 Vilnius, Lietuva.

Tel.: (8 5) 274 4956; faksas (8 5) 270 0112; el. paštas doktor@vgtu.lt

Pranešimai apie numatomą ginti disertaciją išsiųsti 2014 m. lapkričio 17 d.

Disertaciją galima peržiūrėti interneto svetainėje <http://dspace.vgtu.lt/> ir Vilniaus Gedimino technikos universiteto bibliotekoje (Saulėtekio al. 14, LT-10223 Vilnius, Lietuva).

VGTU leidyklos TECHNIKA 2289-M mokslo literatūros knyga

ISBN 978-609-457-734-5

© VGTU leidykla TECHNIKA, 2014

© Lina Papšienė, 2014

lina@rytai.lt

VILNIUS GEDIMINAS TECHNICAL UNIVERSITY

Lina PAPŠIENĖ

THE IMPROVEMENT OF MODELLING
AND HARMONISATION OF SPATIAL
INFORMATION IN DIGITAL CARTOGRAPHY

DOCTORAL DISSERTATION

TECHNOLOGICAL SCIENCES,
MEASUREMENT ENGINEERING (10T)



Vilnius LEIDYKLA
TECHNIKA 2014

Doctoral dissertation was prepared at Vilnius Gediminas Technical University in 2010–2014.

Supervisor

Prof Dr Eimuntas Kazimieras PARŠELIŪNAS (Vilnius Gediminas Technical University, Measurement Engineering – 10T), (2012–2014),
Prof Habil Dr Vytautas GINIOTIS (Vilnius Gediminas Technical University, Measurement Engineering – 10T), (2010–2012).

The Dissertation Defense Council of Scientific Field of Measurement Engineering of Vilnius Gediminas Technical University:

Chairman

Prof Dr Habil Vladas VEKTERIS (Vilnius Gediminas Technical University, Measurement Engineering – 10T).

Members:

Assoc Prof Dr Domantas BRUČAS (Vilnius Gediminas Technical University, Measurement Engineering – 10T),
Prof Dr Marija BURINSKIENĖ (Vilnius Gediminas Technical University, Civil Engineering – 02T),
Assoc Prof Dr Egidijus RIMKUS (Vilnius University, Physical Geography – 06P),
Dr Habil Pawel Aleksander WIELGOSZ (University of Warmia and Mazury in Olsztyn, Measurement Engineering – 10T).

The dissertation will be defended at the public meeting of the Dissertation Defense Council of Measurement Engineering in the Senate Hall of Vilnius Gediminas Technical University at **2 p. m. on 18 December 2014**.

Address: Saulėtekio al. 11, LT-10223 Vilnius, Lithuania.

Tel.: +370 5 274 4956; fax +370 5 270 0112; e-mail: doktor@vgtu.lt

A notification on the intend defending of the dissertation was send on 17 November 2014.

A copy of the doctoral dissertation is available for review at the Internet website <http://dspace.vgtu.lt/> and at the Library of Vilnius Gediminas Technical University (Saulėtekio al. 14, LT-10223 Vilnius, Lithuania).

Reziumė

Disertacijoje pateikti skaitmeninėje kartografijoje taikomi erdvinės informacijos harmonizavimo tyrimai ir jais remiantis patobulinta metodika, detalizuoti metodai erdvinių duomenų pokyčiams nustatyti ir įvertinti.

Darbo tikslas – patobulinti skaitmeninėje kartografijoje taikomą erdvinių duomenų harmonizavimo procesą ir duomenų modeliavimą.

Disertaciją sudaro įvadas, trys skyriai, bendrosios išvados, naudotos literatūros sąrašas, autorės mokslinių publikacijų disertacijos tema sąrašas ir priedai.

Įvadiniamе skyriuje pagrindžiama tiriamoji problema, apibrėžiamas problemos aktualumas, formuluojamas darbo tikslas bei uždaviniai, aprašomas mokslinis darbo naujumas, pateikiami ginamieji disertacijos teiginiai, pristatomos autorės publikacijos ir pranešimai konferencijose, aprašoma disertacijos struktūra.

Pirmajame skyriuje išnagrinėta mokslinė literatūra susijusi su erdvinių duomenų harmonizavimu ir jo galimybėmis erdvinės informacijos infrastruktūros aplinkoje. Apžvelgti automatizuoti erdvinių duomenų generalizavimo metodai ir erdvinių duomenų pokyčių identifikavimo principai. Suformuluoti probleminiai tyrimų klausimai bei uždaviniai darbo tikslui pasiekti.

Antrajame skyriuje pateikta patobulinta erdvinių duomenų harmonizavimo metodika. Sudaryti generalizavimo algoritmai, skirti specifiniams erdvinių duomenų generalizavimo uždaviniams išspręsti. Išskirti erdvinių duomenų pokyčiai, galintys įtakoti erdvinių duomenų skaitmeninėje kartografijoje automatizuoto harmonizavimo procesą, taikant erdvinių duomenų elementų dalinį generalizavimą. Pateikiami erdvinių duomenų pokyčių nustatymo, matavimo ir vertinimo metodai. Pristatytas harmonizuojamų erdvinių duomenų modeliavimo tobulinimas.

Trečiasis skyrius skirtas disertacijos metu parengtos erdvinių duomenų harmonizavimo metodikos taikymui, atliekant georeferencinių erdvinių duomenų rinkinių masteliu 1:10 000 ir 1:50 000 harmonizavimą. Pateikti atlikto harmonizavimo proceso rezultatai.

Disertacijos tema paskelbti 8 moksliniai straipsniai, iš kurių vienas – tarptautinėse duomenų bazėse referuojamuose periodiniuose žurnaluose; du – tarptautiniuose konferencijų leidiniuose; keturi – kitose duomenų bazėse; vienas – respublikinės konferencijos medžiagoje, taip pat perskaityti 9 pranešimai Lietuvos bei kitų šalių konferencijose.

Abstract

The dissertation presents a research on harmonisation of spatial information applied in digital cartography, based on them improved methodology, detail methods for identification and evaluation the changes in features of spatial data.

The main aim of the thesis is to improve the harmonisation process for spatial data used in digital cartography and data modelling.

The thesis comprises an introduction, three chapters, general conclusions, list of references, list of scientific publications by the author on the subject of the dissertation and annexes.

The introduction provides the background to the research problem, defines the relevance of the problem, formulates aims and objectives, describes scientific novelty of the thesis, presents defended statements of the dissertation, author's publications and conference presentations, and describes the structure of the dissertation.

The first chapter analyses scientific literature on harmonisation of spatial data and its abilities in the environment of spatial data infrastructure. Moreover, it presents an overview of automated generalisation methods for spatial data and identification principles of changes in spatial data. Then, it formulates problematic issues of the research and objectives to achieve the aim of the thesis.

The second chapter presents the improved methodology for harmonisation of spatial data. It further presents generalisation algorithms for resolving tasks of the specific generalisation tasks. Furthermore, the chapter centres on changes in spatial data, which may affect automatic harmonisation of spatial data when incremental generalisation is involved in digital cartography. Finally, it introduces identification and evaluation methods for changes in spatial data.

The third chapter is dedicated to the implementation of the methodology for the harmonisation of spatial data relating to (geo)referential spatial data sets at the scales 1:10 000 and 1:50 000. The results of the harmonisation process are presented.

The subject of the dissertation was presented in eight scientific articles, one of which was published in periodicals referenced in international data bases; two in publications of international conferences; four in other data bases; and one in the proceedings of a national conference; besides, there were made nine presentations at conferences in Lithuania and abroad.

Žymėjimai

Santrumpos

CRC – ciklinės perteklinės informacijos tikrinimas (angl. *cyclic redundancy check*).

EII – erdvinės informacijos infrastruktūra, dar vadinama erdvinių duomenų infrastruktūra ar geografinių duomenų infrastruktūra.

ETL – išgauti, pertvarkyti, įkelti (angl. *extract-transform-load*).

ETRS89 – Europos koordinačių sistema (angl. *European Terrestrial Reference System*).

GDR10LT – Lietuvos Respublikos teritorijos georeferencinių erdvinių duomenų rinkinys M 1:10 000.

GDR50LT – Lietuvos Respublikos teritorijos georeferencinių erdvinių duomenų rinkinys M 1:50 000.

GIS – geografinės informacinės sistemos (angl. *Geographic Information System*).

GML – bendros paskirties erdvinių duomenų struktūrų bei jų turinio aprašomoji kalba (angl. *Geography Mark-up Language*).

INSPIRE – Europos bendrijos erdvinės informacijos infrastruktūra (angl. *Infrastructure for Spatial Information in the European Community*).

LKS-94 – Lietuvos valstybinė koordinačių sistema, priimta 1994 metais.

MD5 – žinutės santraukos algoritmas, plačiai naudojama kriptografijos maišos funkcija su 128 bitų (16 baitų) maišos reikšme (angl. *Message-Digest algorithm 5*).

OGC – Atviros geoerdvinės informacijos konsorciumas (angl. *Open Geospatial Consortium*).

VGDM – Vieningas georeferencinių duomenų modelis.

WFS – Internetinė erdviųjų duomenų publikavimo elektroninė paslauga (angl. *Web Feature Service*).

WMS – Internetinė žemėlapių publikavimo elektroninė paslauga (angl. *Web Map Service*).

WPS – Internetinė erdviųjų duomenų apdorojimo elektroninė paslauga (angl. *Web Processing Service*).

XML – bendros paskirties duomenų struktūrų bei jų turinio aprašomoji kalba (angl. *Extensible Markup Language*).

Sąvokos

Duomenų modelis – duomenų struktūrų (lentelių), jų apimamų informacijos laukų, ryšių tarp jų bei operacijų ar funkcijų informacijai jose valdyti aprašymas.

Erdvinė informacija, arba geografinė informacija yra bet koks informacijos įrašas, turintis sąsają su Žemės paviršiumi arba artima Žemės paviršiui erdve.

Erdviniai duomenys, arba geografiniai duomenys – duomenys, kurie tiesiogiai arba netiesiogiai apibūdina konkrečią vietą.

Erdvinių duomenų elementas – realaus pasaulio objekto, esančio konkrečioje vietoje ar geografinėje vietovėje, abstraktus atvaizdas.

Erdvinių duomenų klasė – rinkinys susistemintų erdviųjų duomenų elementų, kurie yra vaizduojami naudojant tą patį geometrijos tipą (taškas, linija, poligonas), kaupiamos tokios pačios kokybinės ir (ar) kiekybinės charakteristikos, ir kurie yra saugomi toje pačioje koordinačių sistemoje.

Erdvinių duomenų rinkinių teikėjai – valstybės kadastrų, registrų tvarkytojai, valstybės ir savivaldybių institucijos bei kiti asmenys, kuriantys ir (ar) tvarkantys erdviųjų duomenų rinkinius.

Erdvinių duomenų rinkinys – susistemintų ir metodiškai sutvarkytų erdviųjų duomenų visuma.

Erdvinių duomenų specifikacija – detalus erdviųjų duomenų rinkinio aprašymas, naudojamas erdviųjų duomenų rinkiniui sukurti.

Geografinės informacijos sistema (arba **geografinė informacinė sistema**) – kompiuterinė sistema, skirta erdvinių duomenų kaupimui, saugojimui, valdymui, pasiekimui ir pateikimui, transformavimui ir analizei, o taip pat – ataskaitų ir išvestinių rezultatų generavimui.

Georeferenciniai erdviniai duomenys – stabilų gamtinių ir antropogeninių objektų ir geodezinio pagrindo punktų erdviniai duomenys.

Informacijos infrastruktūra – visuma kompiuterių, fizinių kompiuterinių tinklų, programinės įrangos, viešosios ar privačiosios skaitmeninės informacijos bei žinių kaip visa tai naudoti. Infrastruktūros elementų naudojimo žinios apima techninius standartus ir protokolus, organizacines priemones, duomenų politiką ir pan.

Internetinė erdvinių duomenų publikavimo elektroninė paslauga – tai OGC specifikacija apibrėžta erdvinių duomenų publikavimo elektroninė paslauga, kuri vektorinius duomenis transformuoja į reikiamą pavidalą kliento pusėje. Ši paslauga leidžia vartotojui atlikti veiksmus su erdviniais duomenimis saugomais serveryje: gauti ar apklausti erdvinis duomenis pagal nustatytus erdvinis arba neerdvinis apribojimus, sukurti naują, patikslinti ar panaikinti egzistuojantį erdvinių duomenų elementą, atnaujinti jo savybes. Šia paslauga erdviniai duomenys yra naudojami kaip įprastiniai vektoriniai duomenys.

Internetinė erdvinių duomenų peržiūros elektroninė paslauga – tai OGC specifikaciją apibrėžta interneto žemėlapių elektroninė paslauga, kuri leidžia pasiekti žemėlapių sluoksnius juos publikuojančiame interneto serveryje pagal duomenų vartotojo pasirinktus parametrus. Ši elektroninė paslauga nesuteikia galimybės pasiekti originalius duomenis vektoriniu formatu, kuriuos naudojant yra paruošiamas žemėlapis publikavimui per elektroninę paslaugą.

Internetinė erdvinių duomenų apdorojimo elektroninė paslauga – tai OGC specifikacija apibrėžta erdvinių duomenų apdorojimo paslauga, kuri leidžia rasti ir pasiekti erdvinis duomenis saugomus serveryje ir juos naudojant atlikti skaičiavimus ar kitus erdvinių duomenų apdorojimo veiksmus.

Maiša (angl. *hashing*) – tai matematinė transformacija, kuri yra atliekama kai tam tikro ilgio simbolių eilutė verčiama į fiksuoto ilgio reikšmę.

Maišos funkcija (angl. *hash function*) – tai funkcija $f = f(x)$, priskirianti argumentui x pseudoatsitiktinį skaičių, vadinamą maišos kodu (angl. *hash code*) arba maišos reikšme (angl. *hash value*). Tam pačiam argumentui funkcija visada duoda tokį pat rezultatą, todėl ji nėra atsitiktinė.

Sąveikumas – sistemos ar jos komponentų gebėjimas užtikrinti informacijos perkėlimą kitoms sistemoms ar jų komponentams. Semantinis sąveikumas reiškia sistemų gebėjimą tarpusavyje keistis informacija neiškreipiant jos prasmės. Techninis sąveikumas reiškia įvairių technologinių sistemų ir jų

komponentų keitimąsi informacija išvengiant intensyvių duomenų konvertavimo užduočių, informacijos praradimo importo ir (ar) eksporto metu bei kitų kliūčių pasiekti paskirstytus duomenis, atsirandančių naudojant heterogenines (įvairiarūšes) informacijos apdorojimo aplinkas ir duomenis.

Tinklo elektroninės paslaugos – apibrėžiamos kaip rinkiniai protokolų ir standartų, nustatančių duomenų apsikeitimo tarp sistemų ar sistemų dalių (aplikacijų) sukurtų skirtingomis priemonėmis ir veikiančių skirtingose aplinkose, būdus.

Transformavimas – (erdvinių duomenų apdorojimo atžvilgiu) yra duomenų pavidalo pakeitimas, pavyzdžiui geodezinių koordinačių sistemos keitimas, erdvinių duomenų elementų geometrijos keitimas, altitudžių vertimas reljefo modeliu ar horizontalėmis, ir pan.

Turinys

IVADAS	1
Problemos formulavimas	1
Darbo aktualumas	2
Tyrimų objektas	3
Darbo tikslas	3
Darbo uždaviniai	3
Tyrimų metodika	4
Mokslinis darbo naujumas	4
Darbo rezultatų praktinė reikšmė	4
Ginamieji teiginiai	4
Darbo rezultatų aprobavimas	5
Disertacijos struktūra	6
Padėka	6
1. ERDVINIŲ DUOMENŲ HARMONIZAVIMO PROBLEMATIKOS SKAITMENINĖJE KARTOGRAFIJOJE ANALIZĖ	7
1.1. Erdvinių duomenų harmonizavimo problematikos erdvinės informacijos infrastruktūrose apžvalga	8
1.2. Erdvinių duomenų automatizuotų generalizavimo metodų apžvalga	15
1.2.1. Generalizavimo kartografijoje apžvalga	16
1.2.2. Generalizavimo operacijos ir algoritmai	20

1.2.3. Generalizavimo modeliai	23
1.2.4. Generalizavimo internete tyrimų apžvalga	25
1.3. Erdvinių duomenų pokyčių identifikavimo analizė	27
1.4. Pirmojo skyriaus išvados ir disertacijos uždavinių formulavimas	30
2. ERDVINIŲ DUOMENŲ HARMONIZAVIMO METODIKOS TOBULINIMAS ...	33
2.1. Erdvinių duomenų harmonizavimo metodikos bendroji dalis.....	34
2.2. Erdvinių duomenų generalizavimo algoritmai	40
2.2.1. Elementų atranka	42
2.2.2. Elementų eliminavimas ir apjungimas	52
2.2.3. Linijinių ir plotinių elementų suprastinimas	56
2.2.4. Linijinių elementų atšakų eliminavimas	68
2.2.5. Plotinio elemento centrinės linijos kūrimas	69
2.2.6. Linijinių ir plotinių elementų sutapdinimas	71
2.3. Erdvinių duomenų pokyčių nustatymo ir įvertinimo metodai	73
2.3.1. Erdvinių duomenų pokyčių nustatymas cikliniu perteklinės informacijos tikrinimo metodu.....	74
2.3.2. Erdvinių duomenų pokyčių įtakos transformavimo procesui vertinimo metodas.....	84
2.3.3. Erdvinių duomenų plotinių elementų pasikeitimo reikšmingumo vertinimo metodika	91
2.4. Harmonizuojamų erdvinį duomenų modeliavimo tobulinimas	96
2.5. Antrojo skyriaus išvados	98
3. ERDVINIŲ DUOMENŲ HARMONIZAVIMO TECHNOLOGIJŲ KŪRIMAS	101
3.1. Georeferencinių erdvinį duomenų harmonizavimo technologijos kūrimas. 101	
3.2. Harmonizuojamų georeferencinių erdvinį duomenų generalizavimas	109
3.3. Harmonizavimo modelio testavimas	115
3.4. Trečiojo skyriaus išvados	124
BENDROSIOS IŠVADOS	127
LITERATŪRA IR ŠALTINIAI.....	129
AUTORĖS MOKSLINIŲ PUBLIKACIJŲ DISERTACIJOS TEMA SĄRAŠAS.....	145
SUMMARY IN ENGLISH.....	147
PRIEDAI.....	163
A priedas. GDR10LT duomenų modelio ištrauka.....	164
B priedas. GDR50LT duomenų modelio ištrauka	179
C priedas. Modelių įgyvendinimo pavyzdžiai.....	189
D priedas. Bendraautorijų sutikimai teikti publikacijose skelbtą medžiagą daktaro disertacijoje	190
E priedas. Autorės mokslinių publikacijų disertacijos tema kopijos.....	196

Contents

INTRODUCTION	1
Problem formulation	1
Relevance of the thesis.....	2
Research object	3
Aim of the thesis	3
Objectives of the thesis	3
Research methodology	4
Scientific novelty of the thesis	4
Practical value of research findings	4
Defended statements	4
Approval of research findings.....	5
Structure of the thesis.....	6
Acknowledgements	6
1. ANALYSIS OF HARMONISATION PROBLEMS OF SPATIAL DATA IN DIGITAL CARTOGRAPHY	7
1.1. Review of harmonisation problems of spatial data in spatial data infrastructures.....	8
1.2. Review of methods of automatic generalisation of spatial data	15
1.2.1. Review of generalisation in cartography.....	16
1.2.2. The operations and algorithms of generalisation	20
1.2.3. Generalisation models.....	23
1.2.4. Review of researches on generalisation on the web.....	25
1.3. Analysis of identification of changes in spatial data	27
1.4. Conclusions of Chapter One and formulation of thesis objectives.....	30

2. IMPROVEMENT OF METHODOLOGY FOR HARMONISATION OF SPATIAL DATA	33
2.1. General part of the methodology for harmonisation of spatial data	34
2.2. The algorithms of spatial data generalisation.....	40
2.2.1. Selection of features.....	42
2.2.2. Elimination and aggregation of spatial data.....	52
2.2.3. Simplification of linear and polygon features.....	56
2.2.4. Elimination of dangles of linear features	68
2.2.5. Creation of a centre line of a polygon feature.....	69
2.2.6. Snapping of linear and polygon features.....	71
2.3. The methods of identification and evaluation of changes in spatial data	73
2.3.1. Identification of changes in spatial data using the method of the Cyclic Redundancy Check.....	74
2.3.2. The method of evaluation of the impact of changes in spatial data to the process of transformation	84
2.3.3. The method for evaluation of significance of changes in polygon features of spatial data	91
2.4. The improvement of modelling of spatial data for harmonisation purposes ..	96
2.5. Conclusions of Chapter Two.....	98
3. DEVELOPMENT OF TECHNOLOGIES FOR HARMONISATION OF SPATIAL DATA	101
3.1. Development of harmonisation technology for (geo)referential spatial data.....	101
3.2. Generalisation of (geo)referential spatial data being of harmonised.....	109
3.3. Testing of the harmonisation model.....	115
3.4. Conclusions of Chapter Three.....	124
GENERAL CONCLUSIONS	127
REFERENCES	129
LIST OF SCIENTIFIC PUBLICATIONS OF THE AUTHOR ON THE TOPIC OF THE THESIS	145
SUMMARY IN ENGLISH.....	147
ANNEXES.....	163
Annex A. The sample of data model of GDR10LT	164
Annex B. The sample of data model of GDR50LT.....	179
Annex C. Examples of the implementation of the models.....	189
Annex D. Agreements of co-authors to provide published materials in the thesis.....	190
Annex E. Copies of the author's scientific publications on the topic of the thesis.....	196

Ivadas

Problemos formulavimas

Daugelį šimtmečių, norint perteikti ir analizuoti realaus pasaulio gamtinius ir antropogeninius objektus, jie buvo kartografuojami ir vaizduojami popieriniuose žemėlapiuose. Taikant geografinės informacijos sistemas (toliau – GIS), šie objektai pradėti vaizduoti naudojant erdvinius duomenis, prireikus juos skaidant į atskirų temų erdvinį duomenų rinkinius, taip suteikiant geresnes duomenų saugojimo ir tvarkymo galimybes, nei sudarant tradicinius popierinius žemėlapius. Todėl šiais laikais žemėlapiai liko tik vienas iš daugelio minėtų objektų kartografinio pateikimo būdų. Plėtojant GIS technologijas atsirado galimybė kaupti didelius kiekius įvairaus tikslumo erdvinį duomenų, aprašant tiek kokybines, tiek kiekybines objektų savybes.

Priklausomai nuo erdvinį duomenų tematikos ir tikslumo, jiems kurti ir atnaujinti taikomi įvairūs būdai. Pavyzdžiui, georeferenciniams ar topografiniams erdviniams duomenims kurti gali būti naudojami geodeziniai matavimai, fotogrametriniai metodai, stambesnio mastelio erdvinį duomenų generalizavimas ir pan. Tačiau tapatūs skirtingo mastelio erdvinį duomenų rinkiniai gali būti tvarkomi savarankiškai, nepriklausomai vieni nuo kitų, taikant skirtingus duomenų modelius, standartus, o duomenys saugomi skirtingais

duomenų formatais ir dažniausiai atnaujinami skirtingais laiko intervalais. Dėl šių priežasčių erdviniai duomenų rinkiniai gali nederėti tarpusavyje. Ypač tai tampa aktualu norint suderinti skirtingų institucijų ar net skirtingų valstybių tiek to paties, tiek skirtingo kartografinio mastelio erdvinius duomenis, kadangi nesuderinti erdvinių duomenų rinkiniai apsunkina kelių rinkinių naudojimą ir analizavimą vienu metu.

Erdvinių duomenų harmonizavimas – tai procesas, suderinantis tarpusavyje skirtingų erdvinių duomenų rinkinių informaciją. Pastarąjį dešimtmetį įgyvendinant harmonizavimo procedūras daugiausiai dėmesio yra skiriama erdvinių duomenų restruktūrizacijai ir transformacijai harmonizuojant to paties ar panašaus kartografinio mastelio tos pačios tematikos erdvinių duomenų rinkinius, detalai netiriant erdvinių duomenų tarp skirtingų mastelių suderinamumą. Todėl sprendžiant erdvinių ar kartografinių duomenų suderinamumą tarp skirtingo mastelio erdvinių duomenų rinkinių, turėtų būti patobulinta erdvinių duomenų harmonizavimo metodika, kuri papildomai numatytų erdvinių duomenų generalizavimo galimybes, keičiančias erdvinių duomenų elementų tipą, mažinančias jų kiekį ar suprastinančias jų sudėtingumą, praleidžiant neesmines duomenų savybes, bet išlaikant svarbias kartografuojamų objektų charakteristikas.

Darbo aktualumas

Skirtingų erdvinių duomenų rinkinių harmonizavimas yra imlus žmogiškiems resursams. Todėl tapatūs erdviniai duomenys skirtinguose rinkiniuose yra dubliuojami ir retai atnaujinami. Harmonizavimo proceso automatizavimas siejant skirtingo mastelio erdvinių duomenų rinkinius įgyvendintų principą, kad erdviniai duomenys būtų kaupiami vieną kartą ir pačiu tinkamiausiu lygmeniu (Europos Parlamento ir Tarybos direktyva... 2007), bei suteiktų galimybę vartotojui gauti su stambesnio mastelio erdvinių duomenų rinkiniu suderintus smulkesnio mastelio duomenis. Tačiau priklausomai nuo harmonizuojamo duomenų kiekio automatizuotam procesui atlikti gali būti reikalingi dideli techniniai resursai. Todėl svarbu papildomai atsižvelgti į galimybę harmonizuoti tik pakitusius duomenis.

Erdvinės informacijos infrastruktūros (toliau – EII) apima technologines priemones, kurios suteikia galimybę teikti vartotojams ne tik erdvinių duomenų rinkinių ar jų dalies kopijas, ar erdvinių duomenų pagrindu sukurtus interaktyvius žemėlapius, tačiau ir priemones, kurios sudaro galimybę harmonizuoti skirtingus erdvinių duomenų rinkinius ir juos pateikti vartotojams kaip visumą. Todėl EII turi veikti erdvinių duomenų harmonizavimo modeliai,

apimantys visumą technologinių priemonių, skirtų erdviniams duomenims harmonizuoti.

Tyrimų objektas

Erdvinių duomenų harmonizavimo metodikos ir skaitmeninėje kartografijoje taikomi erdvinių duomenų modeliavimo ir generalizavimo metodai. Eksperimento objektas – Lietuvos Respublikos georeferencinių erdvinių duomenų rinkiniai M 1:10 000 (toliau – GDR10LT) ir M 1:50 000 (toliau – GDR50LT).

Darbo tikslas

Darbo tikslas – patobulinti erdvinių duomenų harmonizavimo metodiką ir duomenų modeliavimo procesą, kai turi būti transformuojami duomenys į smulkesnio mastelio erdvinių duomenų rinkinį, įvertinant pradinį duomenų pokyčių svarbą.

Darbo uždaviniai

Darbo tikslui pasiekti darbe reikia spręsti šiuos uždavinius:

1. Įvertinti erdvinių duomenų linijinių ir plotinių elementų suprastinimo ir apibendrinimo algoritmus, taikytinus harmonizuojant erdvinius duomenis vaizduojančius gamtinius ir antropogeninius objektus.
2. Ištirti erdvinių duomenų pokyčių pradiname erdvinių duomenų rinkinyje įtaką erdvinių duomenų transformavimui harmonizavimo procese.
3. Patobulinti plotinių erdvinių duomenų pokyčių nustatymo ir įvertinimo metodus taikomus skaitmeninėje kartografijoje.
4. Patobulinti erdvinių duomenų harmonizavimo metodiką, kurią taikant būtų galima sukurti procedūras taikytinas erdvinės informacijos infrastruktūroje, atsižvelgiant į skaitmeninėje kartografijoje taikomus erdvinių duomenų generalizavimo metodus.

Tyrimų metodika

Darbo rengimo metu remtasi užsienio ir Lietuvos mokslininkų publikacijomis ir tyrimais, naudoti skaitmeninėje kartografijoje naudojamų erdvinių duomenų elementų atrankos, suprastinimo ir apibendrinimo algoritmų teoriniai ir eksperimentiniai tyrimai taikant GIS technologijas ir modeliuojant tapačių erdvinių duomenų elementų kitimą pradinų erdvinių duomenų rinkiniuose; atlikta tyrimų rezultatų palyginamoji analizė.

Mokslinis darbo naujumas

Disertaciniame darbe buvo gauti šie matavimų inžinerijos mokslui nauji rezultatai:

1. Patobulinti erdvinių duomenų elementų pokyčių nustatymo, matavimo ir įtakos transformavimo procesui vertinimo metodai, kurie gali būti taikomi erdviniams duomenims harmonizuoti siejant to paties ir skirtingo kartografinio mastelio erdvinių duomenų rinkinius.
2. Patobulintas erdvinių duomenų plotinių elementų harmonizavimo modelis, atsižvelgiant į topologines sąsajas bei taikomas skirtingas tolerancijos reikšmes ir algoritmus.
3. Patobulinta metodika, skirta erdviniams duomenims harmonizuoti, taikytina erdvinės informacijos infrastruktūros aplinkoje.

Darbo rezultatų praktinė reikšmė

Patobulinta erdvinių duomenų harmonizavimo metodika gali būti taikoma sukuriant ir erdvinės informacijos infrastruktūros aplinkoje įdiegiant technologijas, skirtas harmonizuoti mastelio 1:10 000 ir 1:50 000 erdvinis duomenis, taip pat harmonizuoti kitus stambesnio ir smulkesnio kartografinio mastelio duomenų rinkinius.

Ginamieji teiginiai

1. Pasiūlyta erdvinių duomenų harmonizavimo metodika yra tinkama skirtingo mastelio erdviniams duomenų rinkiniams harmonizuoti.
2. Sudarant erdvinių duomenų harmonizavimo modelį siejant skirtingo kartografinio mastelio duomenų rinkinius būtina sudaryti harmonizuojamų kartografinių objektų prioritetinę seką.

3. Harmonizavimo procese potencialios erdvinių duomenų rinkinio atnaujinimo vietos gali būti nustatomos taikant pasiūlytus erdvinių duomenų pokyčių nustatymo, matavimo ir vertinimo metodus.
4. Pasiūlytomis erdvinių duomenų harmonizavimo procedūromis galima aptikti ir eliminuoti nereikšmingus erdvinių duomenų elementų pokyčius.

Darbo rezultatų apibavimas

Disertacijos tema publikuoti 8 moksliniai straipsniai: vienas – tarptautinėse duomenų bazėse referuojamuose periodiniuose žurnaluose (Papšienė, 2013a); du – tarptautiniuose konferencijų leidiniuose (Papšienė *et al.* 2014, Papšienė 2013b); keturi – kitose patvirtintose duomenų bazėse (Papšienė *et al.* 2012; Papšienė, Papšys 2012, 2011, Beconytė *et al.* 2010); vienas – respublikinės konferencijos medžiagoje (Beconytė *et al.* 2010).

Disertacijoje atliktų tyrimų rezultatai buvo paskelbti devyniose mokslinėse konferencijose ir seminaruose Lietuvoje ir užsienyje:

- 9-oje tarptautinėje konferencijoje „*Environmental Engineering*“ 2014 m. Vilniuje;
- 13-oje tarptautinėje daugiadisciplininėje mokslinėje konferencijoje (angl. *13th International Multidisciplinary Scientific GeoConference SGEM*) 2013 m. Albenoje;
- mokslinėje konferencijoje „Nuo kartografijos iki geografinės informacijos mokslo“ skirta M. K. Radvilos Lietuvos Didžiosios kunigaikštystės žemėlapių 400 metų jubiliejui 2013 m. Vilniuje;
- 15-oje Lietuvos jaunųjų mokslininkų konferencijoje „*Mokslas – Lietuvos ateitis*“ 2012 m. Vilniuje;
- 53-oje Rygos technikos universiteto tarptautinėje konferencijoje 2012 m. Rygoje;
- respublikinėje konferencijoje „*Aviacija 2012*“ 2012, Vilniuje;
- 14-oje Lietuvos jaunųjų mokslininkų konferencijoje „*Mokslas – Lietuvos ateitis*“ 2011 m. Vilniuje;
- respublikinėje mokslo konferencijoje „*Civilinė inžinerija ir geodezija*“ 2010 m. Vilniuje;
- respublikinėje konferencijoje „*Mokslas gamtos mokslų fakultete*“ 2010 m. Vilniuje.

Disertacijos struktūra

Disertaciją sudaro įvadas, trys skyriai, bendrosios išvados, literatūros ir šaltinių bei autorės mokslinių publikacijų disertacijos tema sąrašai, santrauka anglų kalba ir priedai.

Darbo apimtis yra 162 puslapiai, neskaitant priedų, tekste panaudota 21 numeruota formulė, 40 paveikslų ir 30 lentelių. Rašant disertaciją naudotasi 187 literatūros šaltiniais.

Padėka

Nuoširdžiai dėkoju moksliniam vadovui prof. dr. Eimuntui Kazimierui Paršeliūnui už mokslines konsultacijas, vertingus patarimus, optimizmą, kantrybę ir pagalbą rengiant disertaciją. Taip pat dėkoju Vilniaus Gedimino technikos universiteto Geodezijos ir kadastro katedros vedėjui prof. dr. Česlovui Aksamitauskui, katedros darbuotojams, Vilniaus universiteto Kartografijos centro dr. Giedrei Beconytei, VĮ Distancinių tyrimų ir geoinformatikos centro direktoriui Evaldui Rožanskui ir visiems kolegoms už vertingus patarimus ir pagalbą. Dėkoju Nacionalinės žemės tarnybos prie Žemės ūkio ministerijos Geodezijos, erdvinės informacijos ir kartografijos skyriui už galimybę naudotis erdviniais duomenimis ir tyrimų rezultatais. Už palaikymą visų doktorantūros studijų metu ypatingai noriu padėkoti savo artimiesiems.

Erdvinių duomenų harmonizavimo problematikos skaitmeninėje kartografijoje analizė

GIS atskyrė anksčiau naudotą erdvinių duomenų ir jų kartografinio vaizdavimo samplą, ir sudarė galimybes valdyti, apdoroti, naudoti ir analizuoti pačius erdvinius duomenis. GIS technologijų taikymas leido efektyviai kaupti ir tvarkyti didelės apimties įvairios tematikos erdvinius duomenis. Tačiau didelis kiekis nesuderintų erdvinių duomenų rinkinių neužtikrina plataus jų panaudojimo. Todėl erdvinių duomenų harmonizavimas tapo pagrindiniu uždaviniu sprendžiant skirtingų duomenų rinkinių suvienodinimą. Šiame disertacijos skyriuje analizuojami erdvinių duomenų harmonizavimo EII principai ir problematika. Apžvelgiami erdvinių duomenų generalizavimo metodai, taikomi skaitmeninėje kartografijoje. Išanalizuoti erdvinių duomenų pokyčių tyrimai.

Skyriuje nagrinėjama tema paliesta autorės kartu su bendraautoriais paskelbtuose keturiuose moksliniuose straipsniuose: Beconytė *et al.* 2010a, 2010b, Papšienė, Papšys 2011, Papšienė *et al.* 2013.

1.1. Erdvinių duomenų harmonizavimo problematikos erdvinės informacijos infrastruktūrose apžvalga

Dar prieš pradėdant plačiai naudoti kompiuterius, Rhind ir Clark (1988) analizavo tos pačios vietovės skirtinguose žemėlapių nomenklatūriniuose lapuose vaizduojamų žemės paviršiaus objektų skirtumus ir nesuderinamumą. Norint naudoti skirtingus erdvinių duomenų rinkinius dažnai iškyla ta pati duomenų suderinamumo problema, kadangi erdviniai duomenys būna sukaupti taikant skirtingus duomenų modelius ir klasifikatorius objektų savybėms apibūdinti. Pavyzdžiui, kuriant M 1:100 000 Europos valstybių administracinių vienetų erdvinių duomenų rinkinį EuroBoundaryMap (anksčiau vadintu SABE (angl. *Seamless Administrative Boundary for Europe*) ar M 1:250 000 Europos topografinių erdvinių duomenų rinkinį EuroRegionalMap, ar M 1:1 000 000 EuroGlobalMap yra naudojami Europos kartografavimo įstaigų sukurti ir palaikomi skirtingų kartografinių mastelių valstybiniai erdvinių duomenų rinkiniai, kurie tarpusavyje yra visiškai nesuderinti. Todėl kuriant vientisą erdvinių duomenų rinkinį, erdvinių duomenų teikėjai prirėkę turi generalizuoti erdvinis duomenis ir vėliau juos pertvarkyti taikant vieningą duomenų modelį ir klasifikatorius (Delattre 2004; Pammer *et al.* 2010; Smith, Rhind 1999).

Vientisos harmoningos erdvinės informacijos poreikis tapo akivaizdus sprendžiant įvairius globalius uždavinius, kuriuose naudojama erdvinė informacija (Estes *et al.* 1995; Collins, Rhind 1997; Lenczowski 1997). Mokslinėje literatūroje nagrinėjamos įgyvendinimo galimybės ir problematika harmonizuojant erdvinę informaciją apie žemės dangą (Jansen 2005; Jensen *et al.* 2008; Herold *et al.* 2006a, 2006b; McConnelli, Moran 2001; Neumann *et al.* 2007), transporto infrastruktūrą (Sandgren, Ottoson 2002); topografinius ir kitos tematikos duomenis (Afflerbach *et al.* 2004; Butenuth *et al.* 2007; Gosseln, Sester 2004; Refsgaard *et al.* 2005; Stoter *et al.* 2009); taip pat derinant erdvinių duomenų rinkinius tarp valstybių (Donaubauer *et al.* 2006; Gedrange *et al.* 2011; Fichtinger *et al.* 2011; Villa *et al.* 2008; Witschas 2005) ir pan.

Didėjant erdvinių duomenų kiekiui, didėjo jos apsiėitimo ir pasiekiamumo svarba. Prieinamumas ir dalinimasis įvairiais erdviniais duomenimis tapo ypač svarbus sprendžiant politinius, ekonominius ar aplinkosauginius klausimus, taip pat nagrinėjant įvairias technines ar mokslines problemas (Smits, Friis-Christensen 2007; Villa *et al.* 2008; Beconytė *et al.* 2010a, 2010b). Tai paskatino kurti EII. Iki 2000 metų daugiausia dėmesio kuriant EII buvo skiriama erdvinės informacijos surinkimui ir jų pasiekiamumui. Tuo metu EII buvo valdomos ir naudojamos tik valstybiniu institucijų lygmeniu, o erdviniai duomenys dažniausiai buvo perduodami magnetinėse ar optinėse laikmenose.

Įsigalėjus internetui, po 2000 metų, duomenims perduoti tarp veikiančių sistemų plačiai pradėtas naudoti interneto ryšys. Atsirado galimybė sukurti lankstų duomenų mainų (teikimo) tinklą, leidžiantį pasiekti erdvinius duomenis iš bet kurios vietos, kurioje tik yra interneto ryšys. Todėl pradėtos kurti EII, leidžiančios teikti erdvinius duomenis aprašančią informaciją ir pačius erdvinius duomenis iš įvairių šaltinių, nepriklausomai nuo jų struktūros, formato ar saugojimo vietos (Beconytė *et al.* 2010a). Siekiant didesnių erdvinių duomenų panaudojimo galimybių, jų harmonizavimas yra vienas iš svarbiausių EII uždavinių.

Pats duomenų harmonizavimas nėra apibrėžiamas vienareikšmiškai. Iš dalies harmonizavimas yra susijęs su duomenų standartizavimu. Masser (2005), apibrėždamas EII, taip pat užsimena apie erdvinių duomenų standartizavimą, kuris turėtų būti vienas iš svarbiausių EII diegimo veiksnių, kadangi nestandartizavus duomenų gali nebūti prasmės jais keistis, nes vartotojams naudotis nestandartizuotais duomenimis gali būti pakankamai sudėtinga. Tačiau standartizavimas gali būti apibrėžiamas, kaip vieningo standarto, aprašančio konkretų duomenų modelį ir klasifikatorius požymiams apibūdinti, naudojimas kaupiant erdvinius duomenis. Tuo tarpu harmonizavimas, tai suvienodinimas duomenų, kurie yra kaupiami naudojant skirtingus duomenų modelius ir klasifikatorius (MacConnell, Morgan 2001). Erdvinių duomenų standartizavimas tiesiogiai įtakoja jos suvienodinimą jau kuriant erdvinių duomenų rinkinius, nepaisant finansinių ir intelektualinių investicijų, o harmonizavimo proceso naudojimas leidžia toliau įprastai kaupti erdvinių duomenų rinkinius, prireikus pertvarkant juos pagal nustatytas taisykles (Jansen 2005). Harmonizavimo procese nėra būtina pašalinti visų skirtumų tarp skirtingų erdvinių duomenų rinkinių, užtenka tik pašalinti nesuderinamumus (Jansen 2004; Herold *et al.* 2006a). Erdvinės informacijos globalaus naudojimo kontekste, harmonizavimas yra suprantamas, kaip galimybė suderinti įvairius erdvinius duomenis, juos integruojant į vientisą ir vienareikšmiškos informacijos duomenų rinkinį, tenkinantį vartotojo poreikius (de Vries *et al.* 2007a).

2007 m. kovo 14 d. priimta Europos Parlamento ir Tarybos direktyva sukurianti Europos bendrijoje EII (toliau – INSPIRE) (Europos parlamento ir Tarybos direktyva... 2007) (toliau – INSPIRE direktyva), apibrėžė 32-i erdvinių duomenų temas, kurių suderinti erdviniai duomenys turi būti teikiami vartotojams per INSPIRE. Todėl vienas iš pagrindinių INSPIRE uždavinių yra užtikrinti teikiamų erdvinių duomenų rinkinių kokybę, sąveikumą ir, prireikus, erdvinių duomenų rinkinių harmonizavimą Europoje (INSPIRE Generic... 2014). EII apimtyje sąveikumas, tai galimybė be daugkartinio žmogaus įsikišimo sujungti erdvinių duomenų rinkinius ir tarpusavyje suderinti juos taip, kad būtų gaunamas darnus rezultatas (Europos Parlamento ir Tarybos direktyva... 2007, Geodezijos ir kartografijos įstatymas 2010), o tam pasiekti

gali būti papildomai atliekamas duomenų harmonizavimas, kuris suteikia prieigą prie erdvinių duomenų per tinklo paslaugas, kurios leidžia nuosekliai sujungti erdvinius duomenis su kitais harmonizuotais duomenimis, naudojant bendrą vieningą duomenų modelį. Todėl erdvinių duomenų harmonizavimą apibendrintai būtų galima įvardinti, kaip procesą suteikiantį prieigą prie erdvinių duomenų rinkinių ir transformuojantį duomenis naudojant vieningą duomenų modelį ir klasifikatorius, tam, kad vartotojai galėtų naudotis sąveikių duomenų visuma nepriklausomai nuo pirminio duomenų šaltinio specifiškumo.

Erdvinių duomenų harmonizavimo problematika ir realizavimas EII aplinkoje yra nagrinėjamas keleto mokslininkų (Giger *et al.* 2014; Hemmatnia *et al.* 2010; Higer *et al.* 2013; Hobona *et al.* 2009; Lehto, Sarjakovski 2004; Reitz, Kuijper 2009; Reitz, Templer 2012; Villa *et al.* 2008; Walters *et al.* 2011). Hintz (2012) išskiria penkis pagrindinius bendruosius etapus harmonizavimo procese: erdvinių duomenų įvertinimas, surinkimas, transformavimas, patikrinimas ir teikimas. Pirmajame etape svarbu įvertinti harmonizuojamų erdvinių duomenų struktūrą, naudojamus klasifikatorius ir suderinamumo galimybes, antrajame – užtikrinti erdvinių duomenų surinkimą iš įvairių duomenų šaltinių, trečiajame – atliekamas erdvinių duomenų transformavimas ir (ar) erdvinių objektų geometrijos keitimas, ketvirtajame – įvertinama gautų rezultatų kokybė, o penktajame etape harmonizuoti erdviniai duomenys teikiami naudotojams.

INSPIRE apimtyje yra išskirti 20 komponentų, kurie įtakoja harmonizavimo procesą. Tai komponentai susiję su bendraisiais principais, terminologija, referenciniu duomenų modeliu, taisyklėmis, skirtomis duomenų struktūroms ir elementų katalogams sudaryti, taip pat susiję su erdviniais ir laiko aspektais, daugiakalbystės adaptavimu, koordinatų ir matavimo sistemomis, objektų referenciniu modeliu, duomenų vertinimo modeliu, pateikties modeliu, identifikatorių valdymu, registrais, metaduomenimis, priežiūra, duomenų ir informacijos kokybe, duomenų perkėlimu ir sutapdinimu, erdvinių duomenų daugialypiu vaizdavimu, duomenų surinkimo taisyklėmis ir erdvinių duomenų atitiktimi aprašytoms taisyklėms (INSPIRE Generic... 2014).

Atitinkamai pastaruoju metu EII aplinkoje kuriami techniniai sprendimai skirti harmonizavimo procese pašalinti skirtumus tarp erdvinių duomenų rinkinių (de Vries *et al.* 2007a, 2007b):

1. Duomenų formato – erdviniai duomenys gali būti saugomi kaip vektoriniai arba rastriniai duomenys, be to skirtingos GIS programinės įrangos gamintojai palaiko skirtingus duomenų formatus. Pavyzdžiui ESRI Shapefile, DWG, MIF ir pan.
2. Koordinatų sistemos – įvairiose šalyse erdviniai duomenys paprastai kaupiami naudojant skirtingą koordinatų sistemą. Pavyzdžiui, Lietuvoje erdviniai duomenys turi būti saugomi valstybinėje Lietuvos

koordinacijų sistemoje LKS-94 (Įsakymas dėl valstybinės koordinacijų sistemos... 1996), INSPIRE apimtyje – Europos koordinacijų sistemoje ETRS89 (INSPIRE Data Specification on Coordinate... 2014).

3. Duomenų modelio – erdviniams duomenims kaupti gali būti naudojami skirtingi duomenų modeliavimo metodai ar duomenų struktūra.
4. Klasifikatorių – tapatiems erdviniams objektams klasifikuoti gali būti naudojami skirtingi klasifikatoriai.
5. Mastelio – skirtinguose rinkiniuose erdviniai duomenys gali būti kaupiami skirtingu tikslumu, priklausančiu nuo taikomo jiems kartografinio mastelio.
6. Detalumo – skirtinguose erdvinių duomenų rinkiniuose gali būti kaupiamas skirtingas kiekis erdvinės informacijos, ar objektus apibūdinančių požymių.
7. Metaduomenų – įvairiose šalyse gali būti naudojami skirtingi metaduomenų standartai erdviniams duomenų rinkiniams aprašyti.
8. Terminologijos – objektai skirtinguose erdvinių duomenų rinkiniuose gali būti semantiškai nevienodai apibūdinami arba naudojama skirtinga kalba jiems aprašyti.
9. Atvaizdavimo – atvaizduojant erdvinius duomenis žemėlapiuose gali būti naudojami skirtingi sutartiniai ženklai, arba skirtinguose erdvinių duomenų rinkiniuose tapatūs objektai gali būti vaizduojami naudojant skirtingus geometrinius elementus.
10. Sutapdinimo – erdvinių duomenų elementų kraštinės, pradžia ar pabaiga gali nesutapti derinant tapačius erdvinius duomenis tarp skirtingų erdvinių duomenų rinkinių.
11. Erdvinių duomenų apdorojimo – skirtinguose rinkiniuose tapatiems rezultatams gauti, gali būti taikomi skirtingi erdvinių duomenų apdorojimo parametrai, funkcijos ir algoritmai. Pavyzdžiui automatizuotai kuriant skaitmeninį reljefo modelį ar generalizuojant erdvinius duomenis.

Erdvinių duomenų formatų ir koordinacijų suderinamumo problema EII šiuo metu yra gana paprastai sprendžiama erdvinių duomenų perdavimui naudojant (Beconytė *et al.* 2010a):

- įvairias standartizuotas tinklo paslaugas, kurios palaiko nuolatinę sąveiką tiesioginę prieigą prie erdvinių duomenų;
- ETL programinę įrangą, kurią naudojant iš anksto sukuriami erdvinių duomenų išrinkimo iš duomenų šaltinio, transformavimo ir pateikimo modeliai, užtikrinantys spartesnę automatizuotą vartotojui reikiamų erdvinių duomenų paruošimą. Šie modeliai leidžia transformuoti duomenis į vartotojo pasirinktą duomenų formatą ir koordinacijų sistemą.

INSPIRE apimtyje erdvinių duomenų teikimas vartotojams yra įgyvendinamas per INSPIRE peržiūros, transformavimo ir atsisuntimo paslaugas (Draft Technical Guidance... 2010; Technical Guidance... 2010; 2013a; 2013b). Tačiau jos yra specifikuotos ir įgyvendintos palyginus neseniai, todėl EII, tame tarpe ir Lietuvos EII, erdviniai duomenys gali būti teikiami naudojant tarptautinio OGC konsorciūmo (angl. Open Geospatial Consortium, Inc.) specifikuotas paslaugas (Beconytė *et al.* 2009; 2010a; Kiehle *et al.* 2007; Steiniger, Hunter 2012): OGC WFS (angl. Web Feature Service (Web Feature Service... 2005) ir OGC WMS (angl. Web Map Service (Web Map Service... 2006).

Erdvinių duomenų modelių, klasifikatorių ir kartografinio mastelio nesuderinamumas sprendžiamas duomenų transformavimo metu. Harmonizavimo procese erdviniai duomenys iš įvairių pirminių šaltinių, kurie naudoja skirtingus duomenų modelius, transformuojami ir perkeliama į vieningą galutinį erdvinių duomenų rinkinį. Bendruoju atveju transformavimas apima erdvinių duomenų apjungimą, generalizavimą ir erdvinių elementų sutapdinimą (Drafting Team... 2008). Lehto (2007a; 2007b) detalizavo ir išskyrė pagrindinius erdvinių duomenų transformavimo proceso komponentus:

1. Atranka (filtravimas):
 - 1.1. Erdvinių duomenų rinkinių atranka – atrenkami tik tie erdvinių duomenų rinkiniai, kurių duomenys turės būti transformuojami.
 - 1.2. Erdvinių duomenų elementų atranka – atrenkami tik tie elementai iš atrinkto erdvinių duomenų rinkinio, kurie turės būti transformuojami.
 - 1.3. Elementų savybių atranka – atrenkamos tik tos atrinktų elementų savybės, kurios atitinka galutinį erdvinių duomenų modelį.
2. Pervadinimas: erdvinių duomenų rinkinių, erdvinių duomenų elementų ir (arba) jų savybių reikšmių pervadinimas (pavyzdžiui, pavadinimų išvertimas iš vienos kalbos į kitą).
3. Perklasifikavimas: erdvinių duomenų rinkinio, erdvinių duomenų elementų ir (ar) jų savybių perklasifikavimas, taikant galutiniame erdvinių duomenų modelyje naudojamus klasifikatorius.
4. Apjungimas arba išskaidymas: erdvinių duomenų rinkinio, erdvinių duomenų elementų ir (ar) jų savybių apjungimas arba išskaidymas.
5. Eilės tvarkos keitimas: erdvinių duomenų elementų rinkinyje ar elementų savybių eilės tvarkos pakeitimas.
6. Erdvinių duomenų elemento keitimas: erdvinių duomenų taškinių, linijinių ar plotinių elementų keitimas iš vieno tipą į kitą (pavyzdžiui, plotinis elementas keičiamas tašku arba jį ribojančia linija).
7. Perskirstymas:
 - 7.1. Erdvinių duomenų rinkinys išskaidomas pagal atskirus erdvinių duomenų elementus, arba elementai pagal atskiras jų savybes.

7.2. Erdvinių duomenų rinkinys sudaromas apjungiant atskirus erdvinių duomenų elementus arba elementai apjungiami pagal atskiras jų savybes.

8. Papildymas:

8.1. Erdvinių duomenų elementų savybių papildymas trūkstamomis reikšmėmis, atsižvelgiant į kitas objekto savybes.

8.2. Erdvinių duomenų elementų numatytųjų savybių galutiniame erdvinių duomenų rinkinyje priskyrimas.

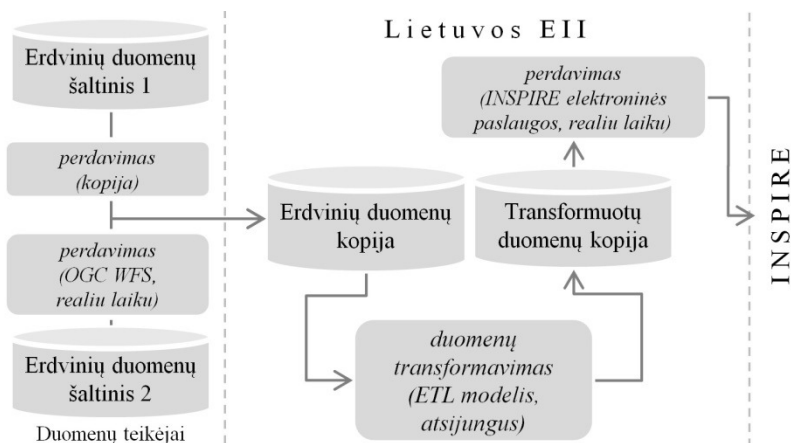
Atsižvelgiant į šiuos komponentus sudaromi erdvinių duomenų transformavimo modeliai, kurie lokaliai arba EII aplinkoje paprastai realizuojami naudojant ETL programinę įrangą arba transformavimo paslaugas (Beconytė *et al.* 2009, 2010a; Hemmatnia *et al.* 2010; Reitz, Templer 2012).

EII įgyvendinant harmonizavimo procesą transformavimas gali būti vykdomas trimis būdais (Lehto 2009, Hemmatnia *et al.* 2010):

1. Transformavimo atsijungus (angl. *off-line*) būdas naudoja iš anksto apibrėžtas transformavimo taisykles ir originalaus šaltinio erdvinių duomenų kopijas. Tačiau šis transformavimo būdas negali užtikrinti transformuojamų erdvinių duomenų aktualumo, kadangi pastarieji gali nuolatos kisti pirminiuose duomenų šaltiniuose.
2. Transformavimas realiu laiku (angl. *on-the-fly*) vykdomas interaktyviai naudojant internetu teikiamas tinklo elektronines paslaugas. Teoriškai toks transformavimo būdas būtų idealus, kadangi nereikėtų saugoti erdvinių duomenų kopijų, nes pats procesas vyktų tiesiogiai naudojant originalaus šaltinio erdvinius duomenis. Tačiau pirmoji problema, su kuria susiduriama naudojant šią erdvinių duomenų transformavimo elektroninę paslaugą, yra greitimeika, kadangi kiekvienu atveju vykdomos mažiausiai dvi transakcijos – geodezinių koordinačių transformavimui ir paties duomenų modelio transformavimui. Antroji problema – kompleksiskumas, kadangi taikomos tiek erdvinės, tiek neerdvinės užklauskos, kurios siejamos su pradinių erdvinių duomenų šaltinio modeliu, kuris paprastai skiriasi nuo duomenų modelio į kurį yra perkeliama transformuoti duomenys.
3. Kombinuotas transformavimas apima abu aukščiau minėtus būdus. Pirmiausiai erdviniai duomenys iš pirminio duomenų šaltinio realiu laiku yra perkeliama ir išsaugomi tarpinėje EII duomenų bazėje, vėliau pats duomenų transformavimas vykdomas lokaliai, po kurio transformuoti duomenys realiu laiku perduodami vartotojams naudojant tinklo paslaugas.

Erdvinių duomenų teikėjai transformavimo būdą pasirenka priklausomai nuo turimos techninės infrastruktūros, programinės įrangos ir esamos EII aplinkos. Pavyzdžiui, Vokietijoje yra taikomas kombinuotas transformavimo

būdas, kai erdviniai duomenys iš pirminių duomenų šaltinių paimami, transformuojami ir išsaugomi tarpinėje duomenų bazėje naudojant ETL, o vėliau realiu laiku teikiama WFS elektroninė paslauga (Hemmatnia *et al.* 2010). Panašus erdvinių duomenų transformavimo būdas yra taikomas ir Lietuvoje įgyvendinant erdvinių duomenų rinkinių harmonizavimą ir teikimą į INSPIRE (1.1 pav.).



1.1 pav. Erdvinių duomenų harmonizavimo principinė schema Lietuvos erdvinės informacijos infrastruktūroje

Fig. 1.1. Common schema of spatial data harmonisation in Lithuanian spatial data infrastructure

Lietuvoje priklausomai nuo erdvinių duomenų teikėjo techninių galimybių, duomenys periodiškai yra perkeltami į Lietuvos EII saugomą erdvinių duomenų kopijų bazę. Prireikus duomenys į ją yra perkeltami naudojant:

- OGC WFS ar kitą tinklo elektroninę paslaugą, teikiančią vektorinius duomenis, kuriuos realiu laiku teikia erdvinių duomenų teikėjas;
- sukurtus ir Lietuvos EII saugomus ETL modelius, kurie prireikus gali tiesiogiai pasiekti ir perkelti iš originalių erdvinių duomenų rinkinių saugojimo vietų;
- erdvinių duomenų teikėjo perduotą ir Lietuvos EII saugomą erdvinių duomenų rinkinio kopiją.

Perkėlus erdvinius duomenis, jie yra transformuojami naudojant iš anksto sukurtus ETL modelius, kuriuose yra įdiegtas originalių erdvinių duomenų transformavimas į INSPIRE duomenų modelius ir ETRS89 koordinatų sistemą. Transformuoti duomenys yra išsaugomi Lietuvos EII transformuotų erdvinių

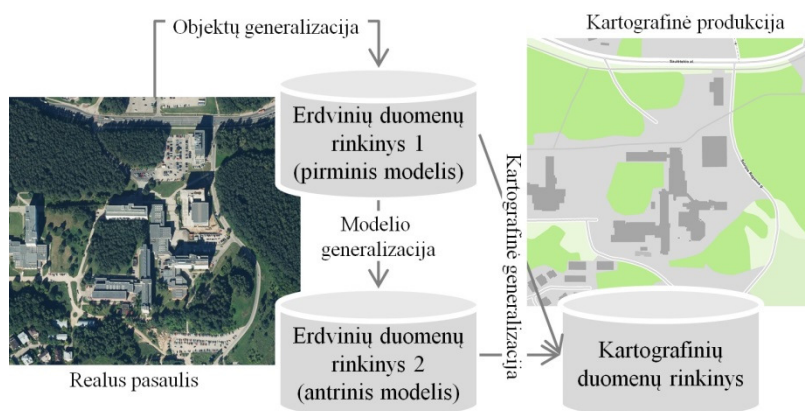
duomenų bazėje. Šie duomenys toliau yra teikiami į INSPIRE realiu laiku naudojant INSPIRE tinklo paslaugas.

Igyvendinant erdvinių duomenų harmonizavimą INSPIRE aplinkoje, INSPIRE erdvinių duomenų specifikacijos nedetalizuoja erdvinių duomenų generalizavimo procesų, tačiau kalba apie tokių procesų įgyvendinimo galimybę erdvinių duomenų transformavimo metu (Drafting Team... 2008; INSPIRE Generic... 2014). Poreikis generalizuoti erdvinius duomenis harmonizavimo procese pirmiausiai gali kilti dėl naudojamų skirtingo erdvinių duomenų detalumo, naudojamų duomenų modelių ar objektų vaizdavimo būdų erdvinių duomenų rinkiniuose. Pavyzdžiui, transformuojant duomenis erdvinių duomenų elementai gali būti apjungiami, eliminuojami ar sukuriama kiti nauji elementai. Tai vieni iš pagrindinių uždavinių vykdant generalizavimą.

1.2. Erdvinių duomenų automatizuotų generalizavimo metodų apžvalga

Procesas, apimantis erdvinių duomenų kiekio mažinimą, pritaikymą konkrečiam masteliui, vadinamas erdvinių duomenų generalizavimu (Müller 1991; Weibel, Dutton 1999). Proceso metu apibendrinama informacija ir sumažinamas jos kiekis, tačiau išlaikomi specifiniai prasmingi kartografinio žemėlapis elementai.

Bendruoju atveju skaitmeninėje kartografijoje yra išskiriamas trejopas generalizavimas: objektų generalizavimas (angl. *object generalisation*), modelio generalizavimas (angl. *model generalisation*) ir kartografinis generalizavimas (angl. *cartographic generalisation*) (Grünreich 1985) (1.2 pav.).



1.2 pav. Generalizavimo seka (pagal Grünreich 1985)
Fig. 1.2. Sequence of generalisation (after Grünreich 1985)

Šių trijų modelių terminologiją pasiūlė Grünreich (1985), kurią vėliau pradėjo naudoti ir kiti mokslininkai.

Erdvinių duomenų rinkinys yra abstraktus realaus pasaulio dalies atvaizdas, todėl erdvinių duomenų kaupimo metu, vektorizuojant erdvinis objektus, jau yra naudojama tam tikro laipsnio generalizavimas, mokslinėje literatūroje vadinamas, kaip objekto generalizavimas. Šio proceso metu kuriamas pradinis erdvinių duomenų rinkinys, vadinamas „pirminiu modeliu“ (Weibel, Dutton 1999).

Modelio generalizavimo, ar kitaip vadinamo erdvinių duomenų generalizavimo metu vykdomas erdvinių duomenų elementų transformavimas, siekiant juos atvaizduoti smulkesnio kartografinio mastelio erdvinių duomenų rinkinyje. Šiuo atveju erdvinių duomenų detalumas ir vaizdavimas suprastinamas tiek, kad jie toliau galėtų būti naudojami įvairiems tikslams. (Weibel, Dutton 1999). Erdvinių duomenų generalizavimas gali būti naudojamas daugiamastelinėse duomenų bazėse, taip pat gali būti taikomas prieš pradendant kartografinį generalizavimą.

Kartografinis generalizavimas, tai žemėlapyje vaizduotinių erdvinių duomenų atranka ir suprastinimas. Skirtingai nuo erdvinių duomenų generalizavimo, atliekant kartografinį generalizavimą erdviniai duomenys yra transformuojami atsižvelgiant į konkrečiam žemėlapiui ar kitai kartografinėi produkcijai keliamus reikalavimus, naudojamus žemėlapyje sutartinius ženklus ir pan. Pavyzdžiui, erdvinių duomenų elementai gali būti perstumiami, perkelti ar specialiai padidinti (Grünreich 1993, Weibel, Dutton 1999).

1.2.1. Generalizavimo kartografijoje apžvalga

Pagrindinis generalizavimo tikslas yra sukurti kokybiškus erdvinis duomenis, išlaikant pusiausvyrą tarp keliamų žemėlapiui tikslumo reikalavimų, kartografinės informacijos apimties ir kartografinio skaitomumo. Todėl generalizavimas apima erdvinės informacijos koregavimą taip, kad ją būtų galima vaizduoti smulkesniu kartografiniu masteliu, tačiau išlaikant esmines kokybines ir kiekybines kartografines žemėlapių charakteristikas. Tarptautinė kartografų asociacija (angl. *International Cartographic Association*) 1973 metais nusakė generalizavimo procesą, kaip „vaizdavimo informacijos atranką ir suprastinimą, atsižvelgiant į mastelį ir (ar) žemėlapių paskirtį“.

Poreikis generalizuoti erdvinis objektus pirmiausiai atsirado kuriant žemėlapius. Pirmieji bandymai inicijuoti automatinį žemėlapių objektų generalizavimą pradėti praeito amžiaus šeštajame dešimtmetyje. Populiarėjant kompiuterinėms technologijoms paplito terminas „skaitmeninė kartografija“ (Shea, McMaster 1989; Buttenfield, McMaster 1991). Erdvinių duomenų ir jų pagrindu kuriamų žemėlapių gamyboje naudojamas modelio (erdvinių

duomenų) ar kartografinis generalizavimas yra sudėtingas daugiakopis uždavinys, priklausantis nuo daugelio veiksnių: pradinių duomenų tikslumo, parinktų metodų, algoritmų ir juose naudojamų parametrų, atliekamų veiksmų eiliškumo, ar sutartinių ženklų naudojamų žemėlapyje. Todėl norint gauti tinkamą rezultatą, turi būti sudarytas tinkamas erdvinų ar kartografinių duomenų generalizavimo modelis, atsižvelgiant į pradinių ir galutinių duomenų mastelius ir jų tematiką.

Erdvinių duomenų informacijos mažinimo procesas jau seniai yra analizuojamas ir aptariamas įvairiose kartografijos, GIS ir kompiuterių mokslų publikacijose. Erdvinių duomenų detalumo mažinimas naudojant generalizavimo metodus gali efektyviai, tačiau kartais radikalai, pakeisti erdvinų objektų savybes: erdvinų duomenų elementų tankumą, plotį, ilgį ar topologines elementų tarpusavio sąsajas. Pagrindinė to priežastis – specifiniai algoritmai, kurių dauguma yra sukurti atsižvelgiant tik į vieną ar kelias minėtas charakteristikas (Cromley, Campbell 1990; Buttenfield 2002), o ne į keliamų reikalavimų visumą. Pastaraisiais dešimtmečiais, atsiradus ir išpopuliarėjus pažangioms informacinėms technologijoms, papildomu generalizavimą įtakančiu faktoriumi tapo erdvinės informacijos perkėlimo tarp skirtingų kartografinių mastelių greitis. Šiandieninių visuomenės poreikių nebetenkina klasikinis vienkartinis erdvinų duomenų generalizavimas. Pavyzdžiui, skaitmeninių interaktyvių internetinių žemėlapių vartotojai tikisi gauti žemėlapių vaizdą ne ilgiau, nei per kelias sekundes. Todėl atsirado poreikis vykdyti erdvinų duomenų generalizavimo procesus realiu ar beveik realiu laiku (Feringa 2001).

Erdvinių duomenų generalizavimas turi du pagrindinius tikslus. Pirma, erdviniai duomenys turi pilnai tenkinti specifinius reikalavimus, antra – pateikiami erdviniai duomenys rinkinyje ar kartografiniai duomenys žemėlapyje turi būti suprantami. To siekiant erdvinų duomenų generalizavimo procesą įtakoja keli veiksniai:

- atrinkti tik esminius erdvinis duomenis ir jų elementus;
- išsaugoti ir pabrėžti būdingus ar neįprastus erdvinų duomenų elementus;
- suprastinti erdvinis duomenis taip, kad tam tikro mastelio žemėlapyje jie būtų suprantami ir įskaitomi.

Šiais pagrindiniais principais intuityviai remiasi specialistai atlikdami erdvinų duomenų generalizavimą rankiniu būdu. Tačiau automatinio generalizavimo atveju turi būti aiškiai įvardinta, ką reiškia svarbūs, būdingi, neįprasti, įskaitomi, suprantami erdviniai duomenys.

Mokslininkas Salichtchev (1976) buvo vienas iš pirmųjų pradėjęs metodiškai analizuoti žemėlapių generalizavimo operacijas, taisykles ir kokybės reikalavimus. 1976 m. mokslininkas Ratajski pristatė vieną iš pirmųjų

konceptinių generalizavimo modelių, kuriame išskyrė du pagrindinius generalizavimo procesų tipus:

- kiekybinė generalizacija, kurioje smulkėjant masteliui palaipsniui mažinamas erdvių duomenų kiekis ir detalumas žemėlapyje;
- kokybinė generalizacija, kurios metu atliekama erdvių objektų transformacija.

Šiame modelyje generalizavimas vykdoma laipsniškai mažinant žemėlapių turinį (kiekybinė generalizacija), tol kol pasiekiamas tinkamas žemėlapių elementų detalumas. Pasiekus tokį lygį, vadinamą „generalizavimo lygiu“ (angl. *generalisation point*) žemėlapių turinio erdvių duomenų kiekis toliau nebegali būti mažinamas neprarandant svarbios informacijos, todėl turi būti atliekama vaizduojamų erdvių objektų transformacija (kokybinė generalizacija), kuria pradedama nauja generalizavimo proceso iteracija. Keletas kitų mokslininkų pasiūlė modelius, skirtus skaitmeniniam generalizavimo procesui (Nickerson, Freeman 1986; Brassel, Weibel 1988; McMaster, Shea 1988; 1992).

1986 metais mokslininkų Nickerson ir Freeman buvo pateiktas generalizavimo procesų modelis, autorės nuomone, yra ribotos apimties. Jų modelio pagrindą sudarė elementų klasifikavimas pagal keturias pagrindines savybes: simbolių pateikimą skirtingais kartografiniais masteliais, elementų savybių perkėlimą, kartografinių mastelių keitimą ir kartografinių pavadinimų išdėstymo optimizavimą. 1988 metais Brassel ir Weibel sukūrė vieną iš labiausiai detalizuotų žemėlapių generalizavimo konceptinių modelių, kuriame aprašomi penki pagrindiniai generalizavimo procesai: erdvių duomenų struktūrinis atpažinimas, procesų atpažinimas, procesų modeliavimas, procesų vykdymas, erdvių duomenų pateikimas.

Pirmieji bandymai siekiant automatizuoti generalizavimo procesą atlikti McMaster ir Shea (1992), kurie papildė anksčiau sukurtąjį Brassel ir Weibel modelį detaliau specifikuodami modelio procesus, bei detalizuodami vykdomas procedūras. Šiame modelyje procesas išskaidytas į tris dalis, siekiant išspręsti šias pagrindines generalizavimo užduotis:

1. Nustatyti priežastis, kodėl tai reikia atlikti.
2. Įvertinti kartometrines sąlygas, identifikuojančias, kada reikia generalizuoti.
3. Parinkti tinkamas erdvių duomenų ir elementų savybių transformacijas, kurios nusakytų, kaip turi būti suprastinami erdvių duomenų elementai.

Tačiau remiantis Müller (1991) pateikta koncepcija, generalizavimas turi patenkinti ne tik erdvių duomenų transformavimo galimybes, siekiant sumažinti erdvių duomenų elementų vaizdavimo kiekį, tačiau turi tenkinti ir keturis pagrindinius reikalavimus susijusius su ekonominiu proceso naudingumu, duomenų aktualumu, daugiatiksliškumu ir vizualizavimu. 1996

metais Ruas ir Plazanet pasiūlė generalizavimo sistemą, kuri būtų kontroliuojama tam tikrų apribojimų. Toks dinaminio generalizavimo modelis pagrįstas apribojimais, susijusiais su erdvinio duomenų elementų grupės savybėmis, būdingomis tam tikrose situacijose: dalyvaujantys generalizavime erdviniai objektai, jų tarpusavio sąsajos ir apribojimai.

1999 m. AGENT (angl. *Automated generalization new technology*) projekto metu buvo sukurtas holistinis generalizavimo modelis, naudojantis taip vadinamas daugiaagentų technologijas. Vietoje to, kad generalizavimo procesas vyktų centralizuotai, atsirado galimybė nustatyti būtinas sąlygas generalizavimui skirtinguose lygmenyse: erdvinio duomenų rinkiniui (mikro-agentai), grupei erdvinio duomenų (meso-agentai) ar erdvinio duomenų elementų klasėms (makro-agentai), tokiu būdu įgalinant juos veikti tiek savarankiškai, tiek komunikuojant su kitais „agentais“. Tokių „agentų“ pagrindinis tikslas yra pagerinti situaciją atsižvelgiant į taikomus apribojimus. Toks naudojamų „agentų“ veiksmų savarankiškumas ir savistovis procesų valdymas sudarė prielaidas sukurti pilnai automatizuotus generalizavimo procesų modelius (Urbanas 2001a, 2001b).

Paskutinio dešimtmečio moksliniuose darbuose pradėtas nagrinėti daugiamastelinių erdvinio duomenų rinkinių palaikymas, interaktyvaus žemėlapių mastelio keitimo galimybės, tiesiogiai susijusios su konkrečiu erdvinio duomenų kartografiniu masteliu ir erdvinio duomenų generalizavimu.

Daugiamasteliniai erdviniai duomenų rinkiniai suteikia galimybę kaupti ir pavaizduoti erdvinio duomenis skirtingais sudėtiniais atvaizdavimais, atsižvelgiant į skirtingiems masteliams taikomus reikalavimus (Kilpelainen 1997; Spaccapietra, Parent. 2000). Tokiuose erdvinio duomenų rinkiniuose erdvinio duomenų atnaujinimai gali būti perduodami kitų mastelių erdvinio duomenims, taip mažinant darbų apimtį tiek žemėlapių, tiek erdvinio duomenų rinkinio informacijai atnaujinti (Bobzien *et al.* 2005; Stoter *et al.* 2004). Tobulėjančios GIS technologijos leido susieti skirtingų kartografinių mastelių erdvinio duomenis tarpusavyje taip sudarant prielaidas vykdyti atnaujinimus erdvinio duomenų elementų lygmenyje, o ne globaliai visų mastelių erdvinio duomenų rinkiniuose (Brewer, Buttenfield 2007). Sprendžiant tokį uždavinį yra išsaugomi erdviniai ryšiai tarp erdvinio duomenų elementų, neatsižvelgiant į kartografinio mastelio pasikeitimus (Tryfona, Egenhofer 1996). Bendrai kalbant visais šiais atvejais visuomet turi būti išlaikoma tiesioginė sąsaja tarp mastelio pokyčio ir generalizavimo lygio (Kreveld 2001). Tokios koncepcijos sukūrimas ypač aktualus internetiniams interaktyviems žemėlapiams, formuojamiems pagal naudotojų poreikavimą (angl. *on-demand*) realiaje laike (angl. *on-the-fly*) (Cecconi, Gallanda 2002; Torun *et al.* 2000;), mobiliems įrenginiams (Harrie *et al.* 2002; Hampe *et al.* 2004), automobilinėje navigacijoje (Li, Ho 2004; Ulugtekin *et al.* 2004) ir kaip papildomos galimybės palaikant daugiamastelinius

erdvinių duomenų rinkinius (Jones *et al.* 2000; Timpf 1998; Zlatanova *et al.* 2004).

1.2.2. Generalizavimo operacijos ir algoritmai

Generalizavimo procesas apima ne tik sprendimą, ką ir kaip generalizuoti, bet ir kaip išspręsti atsiradusius potencialius erdvinių duomenų elementų konfliktus atlikus jų generalizavimą. Naudojant atskiras generalizavimo operacijas yra galimybė sudalinti visą generalizavimo procesą į mažesnės apimties procesus, kurie būtų mažiau komplikuoti. Tokiu būdu pavienes generalizavimo operacijas galima išdėstyti algoritmuose į logišką seką, atsižvelgiant į erdvinių duomenų generalizavimui iškeliamus uždavinius. Išskiriami trys pagrindiniai faktoriai leidžiantys parinkti tinkamiausias operacijas (Cecconi 2003):

- išankstinė situacijos analizė;
- objektų, kuriuos vaizduoja erdviniai duomenys, pobūdis;
- erdvinių duomenų mastelis.

Generalizavimo operacijos nusako abstrakčią erdvinių duomenų elementų transformaciją, o generalizavimo algoritmai apima geometrinę ir semantinę erdvinių duomenų transformaciją.

Per eilę metų vykdant erdvinių duomenų generalizavimo mokslinius tyrimus, generalizavimo operacijos ir terminologija nebuvo standartizuota, todėl skirtingi mokslininkai apibūdindami generalizavimo operacijas, vadovavosi savo tyrimų rezultatais (Beard 1987; Beard, Mackaness 1991; McMaster, Shea 1992; Robinon *et al.* 1984). Pavyzdžiui, McMaster ir Shea (1992) klasifikacijoje išskiriamos operacijos skirtos dviejų tipų transformacijoms (McMaster, Shea, 1992):

1. Erdvinei transformacijai: suprastinimas, glotninimas, agregavimas, amalgamavimas, apjungimas, suglaudininimas, patobulinimas, išdidinimas, perkėlimas.
2. Erdvinių duomenų savybių transformacijai: klasifikacija ir simbolizavimas.

Tačiau reikia pripažinti, kad tiems patiems terminams buvo naudojami skirtingi apibrėžimai, todėl kiti autoriai savo darbuose patikslino juos ir papildė keliomis naujomis operacijomis (Ruas, Langrange 1995). 1999 metais buvo paruoštas naujas generalizavimo operacijų klasifikavimas, kuris buvo panaudotas AGENT projektui (Lamy *et al.* 1999).

Pagrindinės skirtingų mokslininkų išskirtos ir apibendrintos generalizavimo operacijos yra pateikiamos 1.1 lentelėje (Su 1997; Li, Wong 2008).

1.1 lentelė. Mokslininkų išskirtos pagrindinės generalizavimo operacijos**Table 1.1.** The main operations of generalisation distinguished by researchers

Operacijos	Robinson <i>et al.</i> (1984)	Keates (1989)	De Lucia, Black (1987)	McMaster, Monmonior (1989)	Beard, Mackanness (1991)	McMaster, Shea (1992)
Aglomeracija (angl. <i>Agglomeration</i>)			x			
Agregavimas (angl. <i>Aggregation</i>)			x	x		x
Amalgamacija (angl. <i>Amalgamation</i>)				x		x
Apjungimas (angl. <i>Merge</i>)				x		x
Atranka (angl. <i>Selection</i>)		x			x	
Glaudinimas (angl. <i>Collapse</i>)			x	x	x	x
Glotninimas (angl. <i>Smoothing</i>)				x		x
Išdidinimas (angl. <i>Exaggeration</i>)		x			x	x
Išplėtimas (angl. <i>Enhancement</i>)				x		x
Klasifikavimas (angl. <i>Classification</i>)	x				x	x
Kombinavimas (angl. <i>Combination</i>)		x			x	
Pateikimas (angl. <i>Induction</i>)	x					
Patobulinimas (angl. <i>Refinement</i>)			x	x		x
Perstūmimas (angl. <i>Displacement</i>)		x		x	x	x
Praleidimas (angl. <i>Omission</i>)		x		x	x	
Simbolizavimas (angl. <i>Symbolisation</i>)	x					
Suprastinimas (angl. <i>Simplification</i>)	x	x	x	x		x

1.1 lentelės pabaiga

Operacijos	Robinson <i>et al.</i> (1984)	Keates (1989)	Delicia ir Black (1987)	McMaster ir Monmonior (1989)	Beard ir Mackaness (1991)	McMaster ir Shea (1992)
Šiurkštinimas (angl. <i>Coarsen</i>)					x	
Tipifikavimas (angl. <i>Typification</i>)						x

Viena iš pagrindinių erdvių duomenų generalizavimo operacijų yra atranka, kuri yra skirta sumažinti erdvių duomenų ar jų elementų kiekį. Šios operacijos dažnai naudoja elementarias erdvines užklaudas. Jų metu atrenkami pagal nustatytus kriterijus svarbiausi erdviniai duomenys, tokiu būdu pašalinant nesvarbius jų elementus iš duomenų rinkinio. Tokios atrankos dažniausiai grindžiamos kokybiniais ar kiekybiniais kriterijais, objekto svarbumo laipsniu ar elementų tam tikroje teritorijoje tankumu. Jau 1966 metais Töpfer ir Pillewizer pasiūlė būdą, kaip atliekant generalizavimą nustatyti smulkesnio mastelio žemėlapyje vaizduojamų elementų kiekį (tankumą) priklausomai nuo jo ir pradinių duomenų kartografinio mastelio. Tačiau nesant svariems erdvių duomenų atrankos kriterijams reikalingi sudėtingi atrankos algoritmai, kurie analizuotų ne tik erdvių duomenų kiekybines ar kokybines savybes, bet ir elementų geometriją, tarpusavio padėtį ar topologinius ryšius. Erdvių duomenų, vaizduojančių hidrografijos tinklą, elementų atrankai Rusak ir Castner (1990) pasiūlė taikyti algoritmą, kuris leido nustatyti upių segmentų hierarchiją. Thomson ir Richardson (1995) pateikė tinklo analizės prototipą, kuris leido atrinkti kelių tinklo elementus. Pastatus vaizduojančių erdvių duomenų atrankai Ruas (1999) pasiūlė metodą, skirtą atrinkti pastatus atsižvelgiant į jų dydį, tankumą ir kryptį.

Vieni iš pirmųjų matematinių algoritmų pradėti taikyti vykdant linijų suprastinimą (Buttenfield, McMaster 1991; Douglas, Peucker 1973; Richardson 1993; Weibel 1991). Crist (1978) studijavo automatinį taško ir linijos objektų perkėlimą, Lichtner (1979) pasiūlė algoritmą automatiniam erdvių duomenų elemento perkėlimo procesui atsižvelgiant į ryšius tarp elementų, transformuojant erdvinius duomenis iš vieno mastelio į kitą (Lichtner 1979). Ilgą laiką moksliniai tyrimai buvo skirti būtent kuriant erdvių duomenų elementų suprastinimo operacijas. Suprastinimo operacijos pasirinktinai sumažina linijų ar plotinių elementų kontūro viršūnių skaičių. Šiam generalizavimo uždaviniui spręsti yra sukurta daug generalizavimo algoritmų (Visvalingham, Whyatt 1993; Wang, Muller 1998), kurie veikia, kaip nepriklausomų taškų algoritmai, vietinio apdorojimo algoritmai ir kaip globalaus apdorojimo algoritmai (McMaster 1989).

Glottinimo operacijose yra taikomos skirtingos elementų transformacijos, taikant *Fourier* transformaciją (Boutoura 1989; Plazanet *et al.* 1995), išlenkimo transformaciją (Balboa, Lopez 2000), vingiuotumo transformaciją (Burghardt 2005; Burghardt, Meier 1997; Steiniger, Meier 2004).

Daug sukurtų algoritmų yra skirti kitoms linijinių elementų generalizavimo operacijoms atlikti. Pavyzdžiui, dalinei linijos modifikacijai (Bader, Barrault 2000; Li, Su 1996;), kelių tinklo vingių tipifikavimui (Plazanet *et al.* 1995), kelių sankryžų suprastinimui (Mackaness, Mackechnie 1999); viso transporto tinklo generalizavimui (Thomson, Richardson 1995; 1999), lygiagrečių linijų sutraukimui į vieną centrinę liniją (Nickerson 1988); upių tinklo generalizavimui (Rusak-Mazur, Caster 1990), horizontalių generalizavimui (Li, Sui 2000, Ai 2004, Gokgoz 2005) ir pan.

Pagrindiniai algoritmai taikytini plotinių elementų generalizavimui pradėti kurti nuo 1980 metų. Monmonior (1983) sukūrė eilę operacijų skirtų plotinių objektų agregavimui. Vėlesniais metais taip pat buvo sukurta operacijų skirtų ne tik plotinių elementų agregavimui (Su *et al.* 1997b), tačiau ir šių elementų konfigūracijos suprastinimui (Su, Li 1995; Su *et al.* 1997a), perstūmimui (Li, Su 1996), glaudinimui (Su *et al.* 1998). Įgyvendinant generalizavimo operacijas pradėta taikyti trianguliacija ir Voronojaus diagramą plotinių elementų agregavimui ir plotinių elementų centrinių linijų sukūrimui (De Lucia, Black 1987; Jones *et al.* 1995; Li *et al.* 2004).

1.2.3. Generalizavimo modeliai

Pastarąjį dešimtmetį siekiant sumažinti erdvinių duomenų ir jų pagrindu sudaromų žemėlapių paruošimo kaštus bei erdvinių duomenų aktualumą, valstybinės ir privačios kartografavimo įmonės kuria automatizuotas procedūras, kurios galėtų pilnai ar bent dalinai pakeisti rankinį erdvinių duomenų paruošimą. Išskirtinis dėmesys yra skiriamas smulkesnio kartografinio mastelio erdvinių duomenų paruošimui naudojant stambesnio kartografinio mastelio erdvinius duomenis. Galima išskirti kelias priežastis, dėl kurių atsirado poreikis naudoti automatizuotas erdvinių duomenų generalizavimo procedūras. Pirmiausiai tai laiko, kurio reikia atnaujinant rankiniu būdu smulkesnio kartografinio mastelio erdvinius ar kartografinius duomenis, sąnaudos. Pavyzdžiui, Lietuvoje GDR10LT erdviniai duomenys visoje teritorijoje atnaujinami rankiniu būdu per dvejus metus, vektorizuojant naujausiose ortofotografinėse nuotraukose matomą vietovės situaciją. Juos atnaujinus, vykdomas smulkesnių mastelių erdvinių duomenų atnaujinimas. GDR50LT erdvinių duomenų rinkiniui atnaujinti rankiniu būdu prireikia dar papildomai 2–3 metų. Todėl GDR50LT pateikiama informacija gali būti penkerių metų senumo (Papšienė, Papšys 2012, Papšienė 2013b, Papšienė *et al.* 2014). Antroji priežastis – specialistų, atliekančių rankinį

erdvinių duomenų generalizavimą, kvalifikacija ir subjektyvumas. Dėl šios priežasties skirtingų specialistų gaunami rezultatai tikėtina bus taip pat skirtingi (Kilpelainen 2000). Trečioji priežastis yra erdvinių duomenų vietos nustatymo tikslumas, kuris gaunamas blogesnis, kai erdvinių duomenų generalizavimas atliekamas rankiniu būdu, lyginant su kompiuterizuotais automatizuotais metodais (McHaffie 2002). Šios priežastys parodo akivaizdų poreikį siekti, kad erdviniai duomenys ir jų pagrindų sukurti žemėlapiai būtų aktualūs (t. y. atnaujinus, jie vis dar atitiktų realią situaciją vietovėje), standartizuoti ir tikslūs.

Paskutinių metų moksliniai tyrimai generalizavimo srityje leido sukurti programinius sprendimus ir sistemas, skirtas tiek erdvinių duomenų, tiek kartografinių duomenų kūrimui ar atnaujinimui taikant stambesnio mastelio erdvinių duomenų generalizavimą. Pirmosios komercinės programinės įrangos, pasiūliusios specializuotus generalizavimo algoritmus buvo Hanoverio universiteto mokslininkų sukurtas *CPT* (angl. *Change, Push, Tipify*), kompanijos „1Spatial“ sukurtas programinis paketas *Radius Clarity* ir kompanijos „Axpand“ – *Axas System*. Vėlesniais metais galimybė generalizuoti erdvinius duomenis atsirado programinėse įrangoje *ArcGIS* („ESRI“), *Lamps2* („1Spatial“), *DynaGEN* („Intergraph“), *FME* („Safe Software“) ir kt.

2007 metais buvo vykdomas projektas, kurio tikslas buvo įvertinti komercinių programinių įrangų siūlomas automatizuoto generalizavimo galimybės. Šio projekto metu buvo ištestuotos ir įvertintos *CPT*, *Clarity*, *Axpand* ir *ArcGIS* galimybės. Testavimo rezultatai parodė, kad nei viena programinė įranga negali pateikti išbaigto sprendimo, o generalizavimo algoritmus reikia adaptuoti priklausomai nuo erdvinių duomenų modelių (Stoter 2010). Nepaisant to, minėtų programinių įrangų galimybės yra naudojamos nacionalinėse kartografavimo įmonėse, tačiau kiekviena iš jų kuria dalinai ar pilnai automatizuotas generalizavimo procedūras, priklausomai nuo erdvinių duomenų kaupimo logikos, duomenų modelių ir erdvinių duomenų elementams taikomų reikalavimų.

Ispanijos Katalonijos kartografijos ir geologijos institutas (isp. *Institut Cartogràfic de Catalunya*, *ICGC*) sukūrė procesą, skirtą automatizuotai generalizuoti erdvinių duomenų rinkinio M 1:5 000 elementus, kuriant žemėlapi M 1:10 000 ir erdvinį duomenų rinkinį M 1:25 000 (Baella, Pla 1999, 2003, 2005). Tačiau šis procesas yra tobulinamas siekiant kuo geresnių rezultatų atnaujinant kitų mastelių erdvinių duomenų rinkinius. Prancūzijos Valstybinis geografijos institutas (pranc. *Institut Géographique National*, *IGN*) taip pat jau daugelį metų atlieka tyrimus susijusius su kartografiniu generalizavimu: 1999–2005 metais buvo vykdomas projektas *Carto2001* (Lecordix *et al.* 2005), kurio metu naudojant automatizuotą generalizavimą buvo sukurtas *Topo100* žemėlapis M 1:100 000, naudojant georeferencinių duomenų rinkinį *BDCarto* kartografiniu masteliu 1:50 000 (Braun *et al.* 2007). Danijos geoduomenų agentūra (dan.

Kort&Matrikelstyrelsen, KMS) naudoja automatinį generalizavimą kuriant M 1:50 000 žemėlapius iš nacionalinio skaitmeninio topografinio bazinio žemėlapių erdvinių duomenų Top10DK M 1:10 000 (Nielsen, Meyer 2007). Turkijos kartografavimo agentūra (tr. *Harita Genel Komutanligi, HKG*) 2002–2010 metais vykdė projektą KARTOGEN, kurio metu sukūrė modelį, skirtą žemėlapių M 1:50 000 ir 1:100 000 sukūrimui naudojant duomenis kartografiniu masteliu 1:25 000 (Simav *et al.* 2010). Nuo 2006 Italijoje vykdomas CARGEN projektas, kurio metu automatizuotas GeoDBR M 1:5 000 kartografinis generalizavimas, sukuriantis DB25 kartografiniu masteliu 1:25 000, o vėliau DB50 M 1:50 000 (Savino 2011). Olandijos kadastro įmonė (ol. Kadaster) atlikusi kartografinių erdvinių duomenų M 1:50 000 automatizuoto sukūrimo, naudojant M 1:10 000 erdvinių duomenų automatizuoto generalizavimo procedūras, galimybių tyrimus (Stoter *et al.* 2011) 2013 metais pilnai įgyvendino šį procesą. Tokia projektų gausa parodo, kad vis tik nėra sukurtas unikalus modelis, kuris tenkintų įvairių šalių erdvinių duomenų rinkinių ar žemėlapių sukūrimo poreikius.

1.2.4. Generalizavimo internete tyrimų apžvalga

Išpopuliarėjus internetui pradėtos svarstyti galimybės generalizavimo procesą perkelti į pasaulinį tinklą. Tai atlikus, kai kurių mokslininkų nuomone, būtų efektyviai pasidalinta generalizavimo algoritmais ir naudojant elektronines paslaugas būtų galima realizuoti generalizavimo procesą naudojant vartotojo pateikiamus algoritmus. Tokiam dalinimuisi siūloma naudoti internete veikiančias EII ir atviras OGC pagrindu sukurtas elektronines paslaugas. Tačiau iš kitos pusės nustatyta trūkumų taikant atviras OGC elektronines paslaugas ir išskviečiant WPS paslaugas, kadangi nebuvo numatytos galimybės naudoti viena kitą išskviečiančių elektroninių paslaugų (angl. *Invoke Service*) seką, kuri būtina įgyvendinant generalizavimo algoritmus, susidedančius iš atskirų generalizavimo operacijų (Foster, Stoter 2006; Burghardt *et al.* 2005; Neun *et al.* 2008). Įgyvendinant INSPIRE yra specifikuojama elektroninė paslauga, suteikianti galimybę suaktyvinti kitas erdvinių duomenų elektronines paslaugas (angl. *Invoke Service*) (Europos Parlamento ir Tarybos direktyva... 2007; Lucci, Millot 2009).

Keletas mokslininkų nagrinėjo erdvinių duomenų generalizavimo elektroninių paslaugų sukūrimo galimybes ir veikiančios sistemos architektūrą. Buvo atlikta išsami WPS paslaugos architektūros analizė ir veikiančios elektroninės paslaugos projektavimas. Mokslininkams pavyko naudojant WPS paslaugų serverį ir GRASS (angl. *Geographic Resources Analysis Support System*) GIS, atlikti erdvinių duomenų elementų vaizduojančių kelius ir pastatus generalizavimą. Atliekant šį eksperimentą atskirai generalizavimo algoritmai

nebuvo nagrinėjami, kadangi didžiausias dėmesys skirtas technologijų suderinamumo ir veikimui patikrinti. Veikianti sistema buvo apribota vieninteliu įvesties duomenų GML formatu ir vienu žemėlapiu, taip pat negalėjo naudoti išsaugotų (tarpinių) rezultatų (Bergenheim *et al.* 2009; Edwardes *et al.* 2007).

Taip pat buvo atlikti ir generalizavimo, vykdomo kaip elektroninė paslauga internete, metodologijos darbai ir koncepciniai eksperimentai. Foester (2011) analizavo generalizavimo procesą, atliekamą pagal vartotojo pateiktus parametrus. Darbe detalizuota, kokios turi būti sistemos savybės, kad ji galėtų veikti atsižvelgdama į vartotojo pateiktus parametrus, skirtus paruošti žemėlapių beveik realiu laiku. DURP projektas buvo skirtas automatiniam planavimo žemėlapių paruošimui. Tačiau projekto metu buvo prieita prie išvados, kad sukurta sistema gali būti panaudota kaip planavimo portalų koncepcija. Kaip šios sistemos veikimo sąlyga buvo pažymėta, kad jos vartotojas turi būti ypatingai kvalifikuotas ir turi mokėti paruošti sistemai tinkamus parametrus, reikalingus generalizavimui atlikti. Pagal autorių prognozes greitaveiką sistemai galėtų užtikrinti GRID kompiuterijos technologija (Foester *et al.* 2011).

Taip pat buvo ištirtos galimybės atlikti įvairių informacinių sistemų integraciją ir sintezę. Todėl atsirado galimybė projektuoti tokių sistemų architektūrą, kurios geba veikti internete, generuojant įvairaus mastelio žemėlapius, naudodamas įvairų generalizavimui būdingą duomenų apdorojimą. Foerster (2009) analizavo, kokios generalizavimo operacijos būtinos atskirai kiekvienam erdvinio objekto tipui, keičiant vieną mastelį kitu. Pavyzdžiui, generalizuojant pastatus esant kartografiniam masteliui 1:10 000 ir pateikiant juos kartografiniu masteliu 1:50 000, amalgamacija (angl. *amalgamation*) yra labai svarbi, o suglaudavimo (angl. *colappse*) ir prastinimo (angl. *simplification*) operacijos yra vidutiniškai svarbios, o likusios generalizavimo operacijos praktiškai nėra svarbios.

Šioje srityje atlikti ir kiek kitokios pakraipos moksliniai darbai, kuriais buvo siekta išsiaiškinti generalizavimo taikymo galimybes ruošiant žemėlapius internete. Tačiau Zhou su kolegomis (1999) susidūrė su žemėlapių ruošimo greitaveikos problema, todėl buvo prieita išvados, kad geriau yra iš anksto paruošti erdvinį duomenų rinkinius generalizavimui. Tokiu būdu generalizavimas gali būti atliekamas naudojant gana paprastas technologijas, norint sukurti patenkinamos vizualios kokybės suprastintą žemėlapi.

Generalizavimo problematika mokslo darbuose yra išskaidyta į atskirus komponentus, bet neverta užmiršti ir sąsajos tarp kartografinio generalizavimo ir šiuolaikinės daugiamastelinės (angl. *multi-scale*) duomenų bazės palaikymo problematikos. Šios sąsajos apibūdinimas ir vieta yra svarbi tolimesniems tyrimams, todėl ji gana dažnai nagrinėjama įvairiais aspektais. Cecconi (2003) pateikia keletą idėjų – jo nuomone generalizavimui atlikti būtinos daugiamastelinės duomenų bazės ir jos turi būti praturtintos papildomais

duomenimis, kurie leistų optimaliau atlikti generalizavimą. Be to realizuojant generalizavimo procesą realiu ar beveik realiu laiku (angl. *on-the-fly*), siūloma ignoruoti tas generalizavimo operacijas, kurios reikalauja daug laiko sąnaudų.

1.3. Erdvinių duomenų pokyčių identifikavimo analizė

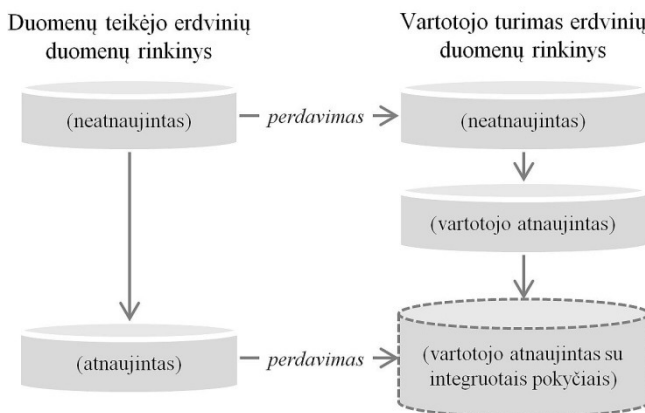
Šiuolaikinės GIS technologijos leidžia plačiai naudoti erdvinius duomenis, juos integruoti į kitus erdvinius duomenų rinkinius, analizuoti ir plėsti pagal poreikius. Todėl erdvinių duomenų pokyčių nustatymo procesas yra svarbus harmonizavimo procese siekiant ne tik sumažinti sąnaudas, reikalingas erdvinių duomenų atnaujinimo procesui, bet ir išlaikant informaciją apie nepakitusius erdvinių duomenų elementus ar sekant jų gyvavimo ciklą.

Erdvinių duomenų pokyčių suradimas yra procesas identifikuojantis skirtingą erdvinio objekto būseną analizuojamu laikotarpiu (Singh 1989). Identifikuoti erdvinių duomenų pokyčiai gali būti naudojami aplinkai vertinti (pavyzdžiui, žemės dangos pokyčių stebėjimui) arba siekiant identifikuotus pokyčius perduoti susijusiam erdvinių duomenų rinkiniui.

Pokyčių paieška vykdoma tiek rastriniuose, tiek vektoriniuose erdvinių duomenų rinkiniuose. Tačiau pokyčių nustatymo metodai buvo pirmiausiai pradėti plačiai taikyti rastriniuose vaizduose identifikuojant ir išskiriant žemės dangos pokyčius (Feranec *et al.* 2007; Lu *et al.* 2004; Lunetta *et al.* 2006; Mas 1999), taip pat vykdant miškų stebėseną (Desclee *et al.* 2006; Hame *et al.* 1998) ir panašiose srityse. Pokyčio tarp skirtingų rastrinių vaizdų, tokių kaip kosminės ar ortofotografinės nuotraukos, atpažinimo algoritmai yra plačiai nagrinėjami mokslinėje literatūroje (Singh 1989; Mas 1999; Lu *et al.* 2004; Radke *et al.* 2005; Volpi *et al.* 2011, Milieškaitė 2010, 2011, Milieškaitė, Vaitkus 2011).

Poreikį nustatyti pokyčius erdvinių duomenų rinkinyje pirmiausiai įtakojo duomenų naudotojai, kurie duomenų teikėjo teikiamą erdvinių duomenų rinkinį naudoja atnaujinant savo palaikomus erdvinius duomenis (Peerbocus *et al.* 2002), kuriuose gali atsirasti pokyčių per tą laikotarpį, kol bus gautas atnaujintas duomenų teikėjo erdvinių duomenų rinkinys (1.3 pav.). Pavyzdžiui, atnaujinus georeferencinio pagrindo objektų erdvinius duomenis, teminių erdvinių duomenų rinkinyje gali kisti erdviniai duomenys vaizduojantys minėto objekto apsauginę zoną (Beconytė *et al.* 2010a).

Erdvinių objektų pokyčių nustatymo tarp dviejų erdvinių duomenų rinkinių uždavinys pradėtas nagrinėti praeito tūkstantmečio pabaigoje. Pirmieji siūlomi sprendimai rėmėsi erdvinių objektų konfigūracijos palyginimu (Devogele 1998).



1.3 pav. Pokyčių integravimas (Peerbocus *et al.* 2002)

Fig. 1.3. Integration of updates (Peerbocus *et al.* 2002)

Dauguma siūlomų automatizuoto pokyčio tarp skirtingų erdvinių duomenų rinkinių nustatymo būdų yra pagrįsti visų erdvinių objektų palyginimu erdviniuose duomenų rinkiniuose prieš ir po atnaujinimo (Badart 1998) taikant erdvinių duomenų elementų sutapdinimo algoritmus (Lemarie *et al.* 1996). Sutapdinimo metu nustatomas ryšys tarp erdvinių duomenų elementų skirtinguose erdvinių duomenų rinkiniuose. Priklausomai nuo nustatytų ryšių kiekybinio tipo (angl. *cardinality*), juos galima išskirti į tris pagrindines grupes (Peerbocus *et al.* 2002):

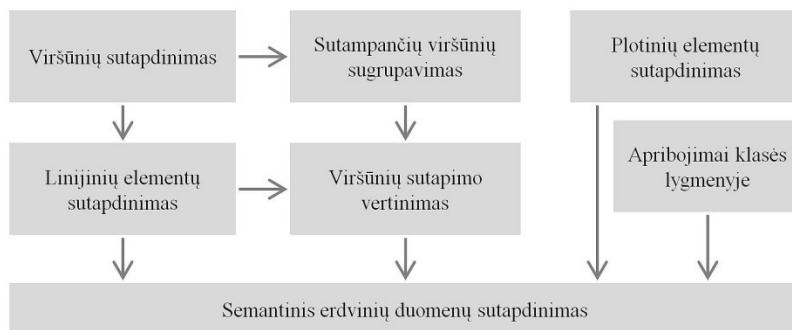
- 1:0 arba 0:1 reiškia, kad elementas neturi atitikmens kitoje erdvinių duomenų rinkinio versijoje;
- 1:N arba N:1, kai $n > 0$ reiškia, kad elementas turi vieną ar kelis atitikmenis kitoje duomenų rinkinio versijoje;
- N:M, kai $N > 1$ ir $M > 1$ nurodo, kad keli elementai erdvinių duomenų rinkinyje atitinka kelis kitus elementus kitoje erdvinių duomenų rinkinio versijoje.

Nustačius tokius ryšius tarp erdvinių duomenų galima analizuoti pokyčius ir juos klasifikuoti priklausomai nuo jų pobūdžio (Bahard 1998). Iškiriami dviejų tipų erdvinių duomenų pokyčiai: erdviniai, kai pasikeičia erdvinio objekto konfigūracija arba topologija tarp kaimyninių erdvinių objektų, ir savybių pokyčiai, kai keičiasi erdvinio objekto savybės. Duomenų pokyčio nustatymas, kai yra lyginami du duomenų rinkiniai, apima keturis pagrindinius procesus (Armeniakis *et al.* 2002):

- pokyčio radimas,
- atpažinimas, nustatant pokyčio pobūdį,
- identifikavimas,

- kiekybinis įvertinimas, skirtas įvertinti pokyčio svarbą.

Mokslinėje literatūroje siūlomi skirtingi sutapdinimo (ryšio tarp erdviųjų objektų nustatymo) algoritmai priklausantys nuo erdviųjų objektų tipo ir sudėtingumo (Badard, Lemarie 1999; Gosseln, Sester 2004; Harrie, Hellstrom 1999; Li, Goodchild 2010; 2011; Safra *et al.* 2010; Yuhong *et al.* 2013; Walter, Fritsch 1999). Badard (1999) sukūrė algoritmą, skirtą pokyčių erdviame duomenų rinkinyje paieškai, naudodamas erdviųjų duomenų elementų, esančių skirtinguose erdviųjų duomenų rinkiniuose, sutapdinimą (1.4 pav.).



1.4 pav. Bendrinė erdviųjų duomenų sutapdinimo mechanizmo schema (Badard 1999)

Fig. 1.4. General scheme of mechanism of spatial data matching (Badard 1999)

Tačiau sutapdinimo metu gali iškilti semantinio nesuderinamumo problemos (Comber *et al.* 2004). Gosseln ir Sester (2004) pasiūlė sprendimą automatizuotam pokyčių nustatymui ir erdviųjų duomenų rinkinio atnaujinimui. Jų siūlomas sprendimas apėmė suskaidytų erdviųjų duomenų elementų apjungimą, sąsajų tarp tapačių elementų nustatymą ir sankirtos operacijos panaudojimą, norint nustatyti pokyčius.

Bendruoju atveju pokyčio radimas apima galimybę atlikti kokybinę jo analizę, įvertinant galimą poveikį erdviams duomenims (Singh 1986, Lu *et al.* 2004). Tam tikslui pasiekti turi būti analizuojamas erdviųjų duomenų atvaizdo galimas kitimas keičiant įvairius parametrus ir išskiriant veiksmus įtakančius ir neįtakančius vaizduojamų elementų kitimą (Green *et al.* 1994). Tai yra pagrindinis uždavinys siekiant pašalinti nereikšmingus pokyčius ir atrenkant tik tuos, kurie turės įtaką erdviųjų duomenų kitimui (Radke *et al.* 2005).

Dauguma sukurtų pokyčių nustatymo algoritmų yra skirti panašaus kartografinio mastelio erdviųjų duomenų rinkiniams analizuoti, todėl jie nėra tinkami norint nustatyti pokyčius stambesnio mastelio erdviuose duomenų rinkiniuose, siekiant įvertinti jų įtaką smulkesnio mastelio erdviųjų duomenų rinkiniui.

1.4. Pirmojo skyriaus išvados ir disertacijos uždavinių formulavimas

Atlikus mokslinės literatūros analizę galima daryti tokias išvadas:

1. Pastaruoju metu erdvinių duomenų harmonizavimas EII aplinkoje yra orientuotas į erdvinių duomenų rinkinių transformavimą naudojant vieningą duomenų modelį, siekiant, kad skirtingų šaltinių erdviniai duomenys būtų nuosekliai sujungti ir juos būtų galima naudoti kaip vieną visumą.
2. Atlikus mokslinės literatūros analizę, nustatyta, kad nepakankamai išsamiai išanalizuotos galimybės naudoti erdvinių duomenų generalizavimą atliekant erdvinių duomenų harmonizavimą kombinuotu būdu, kai erdvinių duomenų transformavimas vykdomas atsijungus (angl. *off-line*).
3. Esamos erdvinių duomenų generalizavimo sistemos skaitmeninėje kartografijoje siūlo eilę skirtingų sprendimų, tačiau pats tinkamiausias sprendimas nėra rastas, kadangi visais atvejais analizuojant modelius ir algoritmus paaiškėja, kad siūlomas procesas turi būti pritaikytas konkrečiam erdvinių duomenų modeliui.
4. Pavienių generalizavimo algoritmų taikymas išsprendžia tik erdvinių duomenų elementų vaizdo suprastinimą, tačiau neatsižvelgia į objektų tarpusavio sąsajas, todėl reikalinga sukurti modelius, apimančius logiškai išdėliotų algoritmų seką, kuri apimtų ne tik pavienių objektų generalizavimą, bet ir su jais susijusių objektų pokyčius, topologinių sąsajų atstatymus.
5. Erdvinių duomenų generalizavimas internete naudojant elektronines tinklo paslaugas yra naudojamas, tačiau šios paslaugos neapima sudėtingų erdvinių duomenų generalizavimo procedūrų.

Atsižvelgiant į literatūros analizės rezultatus, išryškintus probleminius klausimus bei darbo tikslą, disertaciniame darbe reikalinga išspręsti tokius pagrindinius uždavinius:

1. Atrinkti, ištirti ir įvertinti erdvinių duomenų linijinių ir plotinių elementų suprastinimo ir apibendrinimo algoritmus, kurie galėtų būti naudojami harmonizavimo procese transformuojant erdvinius duomenis vaizduojančius gamtinius ir antropogeninius objektus.
2. Ištirti erdvinių duomenų pokyčius pradinių erdvinių duomenų rinkinyje ir jų sąsają su erdvinių duomenų elementais atlikus jų transformavimą ir generalizavimą.
3. Siekiant išmatuoti elemento pokyčius ir nustatyti jų reikšmingumą, patobulinti erdvinių duomenų pokyčių nustatymo ir įvertinimo metodą taikomą skaitmeninėje kartografijoje.

4. Patobulinti erdvinių duomenų harmonizavimo metodiką, kurią taikant būtų galima sukurti procedūras taikomas erdvinės informacijos infrastruktūroje, atsižvelgiant į skaitmeninėje kartografijoje taikomus erdvinių duomenų generalizavimo metodus, ir atlikti jų tyrimus.

Erdvinių duomenų harmonizavimo metodikos tobulinimas

Antrajame skyriuje pateikiama patobulinta erdvinių duomenų harmonizavimo metodika, kurią galima taikyti kuriant skirtingų kartografinių mastelių erdvinių duomenų rinkinių harmonizavimo technologiją. Vykdamas tokių erdvinių duomenų rinkinių harmonizavimą būtina taikyti erdvinių duomenų elementų generalizavimą. Todėl ištirti pagrindiniai generalizavimo algoritmai taikomi linijiniams ir plotiniams elementams, taip pat atlikti linijinių elementų suprastinimo algoritmų teoriniai tyrimai. Atliktas eksperimentas, lyginant Douglas-Peucker (Douglas, Peucker 1973), Wang (Wang, Muller 1999), ir siūlomą praplėstą Wang linijų suprastinimo algoritmus. Taip pat sudaryti generalizavimo algoritmai, skirti specifiniams erdvinių duomenų generalizavimo uždaviniams išspręsti, naudojant GIS programinę įrangą. Vadovaujantis gautais rezultatais, patikslinti erdvinių duomenų pokyčių nustatymo, matavimo ir vertinimo metodai. Skyriaus tematika atliktų tyrimų rezultatai paskelbti autorės straipsniuose (Beconytė *et al.* 2010, Papšienė 2013a, 2013b) ir konferencijose (Civilinė inžinerija ir geodezija 2012; 53-oje Rygos technikos universiteto tarptautinėje konferencijoje 2012; 13-oje daugiadisciplininėje tarptautinėje mokslinėje konferencijoje Albenoje, Bulgarija 2013).

2.1. Erdvinių duomenų harmonizavimo metodikos bendroji dalis

Disertacijos darbo tikslas yra patobulinti erdvinių duomenų harmonizavimo metodiką, kurią taikant būtų galima sukurti procedūras EII ir technologijas, atsižvelgiant į galimybę harmonizuoti skirtingo kartografinio mastelio erdvinius duomenis, taikant skaitmeninėje kartografijoje naudojamus erdvinių duomenų generalizavimo metodus.

Siūloma patobulinta metodika yra sukurta atsižvelgiant į RISE (RISE 2007) ir HUMBOLD (Fichtinger *et al.* 2011) projektų metu sukurtas erdvinių duomenų harmonizavimo EII bendrąsias metodikas ir INSPIRE išskirtus harmonizavimo komponentus (INSPIRE Generic... 2014). Siūloma patikslinta metodika praplėsta atsižvelgiant į poreikį harmonizuoti skirtingo kartografinio mastelio erdvinius duomenis. Metodika papildyta erdvinių duomenų pokyčių nustatymo ir vertinimo procedūromis, išskiriant plotinių elementų konfigūracijos pokyčio įvertinimą.

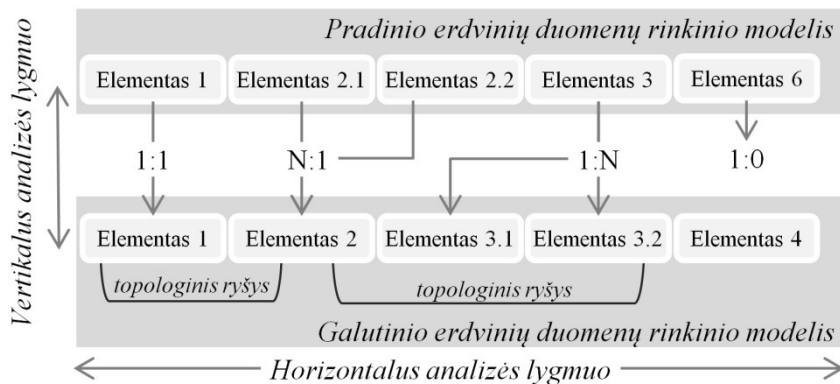
Norint harmonizuoti erdvinius duomenis, būtina pirmiausiai parinkti tinkamus pradinius erdvinių duomenų rinkinius, kurie harmonizavimo metu turi būti transformuojami ir perkeliama į galutinių erdvinių duomenų rinkinį.

Erdvinių duomenų harmonizavimo technologijos kūrimas turi apimti penkis pagrindinius etapus: analizės, projektavimo, kūrimo, testavimo ir dokumentavimo. Erdvinių duomenų harmonizavimo proceso kūrime analizės etapas yra vienas iš svarbiausių, kadangi vadovaujantis gautais rezultatais, kitame etape turi būti sukurtos transformavimo priemonės, kurios neveiks tinkamai, jei bus panaudoti neteisingi rezultatai. Analizės etape turi būti atlikta:

1. Pradinių erdvinių duomenų rinkinio ir galutinio erdvinių duomenų rinkinio analizė, kurios metu nustatoma prieigos galimybės prie duomenų, erdviniams duomenims taikoma geodezinių koordinačių sistema, matavimo vienetų sistema ir erdvinių duomenų saugojimo formatas.
2. Pradinių erdvinių duomenų rinkinio ir galutinio erdvinių duomenų rinkinio modelių analizė, kurios metu analizuojama naudojama terminologija, erdvinių duomenų elementų klasės, juose saugomų elementų geometrijos tipai ir naudojami klasifikatoriai elementų savybėms apibūdinti, erdvinių duomenų rinkinio elementų tikslumas ir detalumas priklausantis nuo kartografinio mastelio.
3. Erdvinių duomenų elementų ir jiems priskiriamų savybių reikšmių analizė, nustatant ryšį tarp tapačių elementų aprašytų skirtinguose erdvinių duomenų modeliuose, atsižvelgiant į elementams priskiriamas savybių reikšmes (vertikalus analizės lygmuo) (2.1 pav.). Priklausomai

nuo nustatytų ryšių kiekybinio tipo (angl. *cardinality*), juos reikia išskaidyti į penkias grupes (pagal Peerbocus *et al.* 2002), kai:

- 3.1. vienas elementas pradiniam erdvinių duomenų modelyje turi vieną atitikmenį galutiniame erdvinių duomenų modelyje (1:1);
- 3.2. vienas elementas pradiniam erdvinių duomenų modelyje turi kelis atitikmenis galutiniame erdvinių duomenų modelyje (1:N, kai $N > 0$);
- 3.3. keli elementai pradiniam erdvinių duomenų modelyje turi vieną atitikmenį galutiniame erdvinių duomenų modelyje (N:1, kai $N > 0$);
- 3.4. elementai neturintys atitikmenų (1:0, 0:1);
- 3.5. keli elementai pradiniam erdvinių duomenų modelyje turi kelis atitikmenis galutiniame erdvinių duomenų modelyje (N:M, kai $N > 1$ ir $M > 1$).



2.1 pav. Erdvinių duomenų elementų sąsajų pavyzdys

Fig. 2.1. Example of relations between elements of spatial data

Ryšio nustatymo principas parodytas 2.1 paveiksle: „Elementas 1“ pradiniam erdviniame duomenų rinkinyje turės atitikmenį galutiniame erdviniame duomenų rinkinyje (ryšys 1:1), „Elementas 2.1“ ir „Elementas 2.2“ pradiniam erdvinių duomenų rinkinyje atitiks vieną „Elementas 2“ galutiniame rinkinyje (ryšys N:1), o „Elementas 3“ atitiks „Elementas 3.1“ ir „Elementas 3.2“ (ryšys 1:N). Tuo tarpu „Elementas 6“ jokio ryšio neturės su kitu elementu, esančiu galutiniame erdvinių duomenų rinkinyje (ryšys 1:0). Ryšys N:M 2.1 paveiksle nėra pavaizduotas, tačiau tai yra bendras ryšių 1:N ir N:1 atvejis.

4. Erdvinių duomenų elementų galutiniame erdvinių duomenų rinkinio modelyje analizė ir topologinių ryšių tarp elementų nustatymas

(horizontalus analizės lygmuo). Kaip pavaizduota 2.1 paveiksle erdvinių duomenų elementai rinkinyje gali turėti tarpusavio ryšius, kurie gali būti tiek tarp tos pačios erdvinių duomenų elementų klasės elementų, tiek tarp skirtingų elementų klasių (2.2 pav.).



2.2 pav. Pavyzdys topologinių ryšių tarp erdvinių duomenų
Fig. 2.2. Example of topologic relations between elements of spatial data

5. Papildomų taisyklių, kurios turi būti taikomos pradinio erdvinių duomenų rinkinio elementams, nustatymas, atsižvelgiant į pradinio ir galutinio erdvinių duomenų rinkinių modelių analizės rezultatus. Pavyzdžiui, norint harmonizavimo procese transformuoti linijinius elementus taip, kad būtų sukurti plotiniai elementai, būtina nustatyti, ar pradiname erdvinių duomenų rinkinyje visų linijinių elementų pradžia ir pabaiga turi bendrą viršūnę su kito linijinio elemento pradžia ar pabaiga.

Erdvinių duomenų harmonizavimo skirtingų etapų rezultatams saugoti siūloma sukurti duomenų bazes, skirtas pradinio erdvinių duomenų perkėlimui (toliau – atnaujinimo šaltinių erdvinių duomenų bazė), pradinio erdvinių duomenų su identifiкуotais pokyčiais saugojimui (toliau – tarpinių erdvinių duomenų bazė), jei harmonizavimo procese svarbu identifiкуoti pokyčius, transformuotų erdvinių duomenų saugojimui (toliau – galutinių erdvinių duomenų bazė), rastų klaidų saugojimui (toliau – klaidų duomenų bazė). Jei

transformavimo procese yra taikomas generalizavimas, tuomet rekomenduojama sukurti atskiras duomenų bazines, skirtas generalizuotiems erdviniais duomenims (toliau – tarpinė transformuotų duomenų bazė). Vėliau šiems duomenims atliekamos likusios transformavimo operacijos ir duomenys perkeliami į galutinių erdvinį duomenų bazę.

Projektuojant harmonizavimo technologiją, ji turėtų apimti ne mažiau, kaip šiuos žemiau išvardintus procesus, užtikrinančius korektišką erdvinį duomenų harmonizavimą, ir skirtus:

1. Pradinių erdvinį duomenų rinkinio pasiekimui ir perkėlimui. Šio proceso metu turės būti inicijuojamas erdvinį duomenų gavimas iš pirminio duomenų šaltinio ir perkėlimas į atnaujinimo šaltinių erdvinį duomenų bazę.
2. Atnaujinimo šaltinių erdvinį duomenų automatizuotai patikrai. Ši dalis nėra aptariama kitose metodikose, tačiau tokios patikros tikslas yra patikrinti pagrindines erdvinį duomenų savybes, dėl kurių neteisingumo gali būti nekorektiškai transformuoti erdviniai duomenys. Ypač tai svarbu tuomet, kai harmonizavime vykdomas transformavimą turi būti taikomi erdvinį duomenų generalizavimo metodai.
3. Erdvinių duomenų perkėlimui į tarpinių erdvinį duomenų bazę. Šio proceso metu turi būti vykdomas erdvinį duomenų pokyčių identifikavimas ir įvertinimas remiantis 2.3 poskyryje aprašytu metodu. Šis procesas turi leisti perkelti ir integruoti tik pakitusius erdvinis duomenis. Nebeegzistuojantys erdviniai duomenys ištrinami iš tarpinės duomenų bazės. Šis procesas turi būti vykdomas tuomet, kai atlikus harmonizavimą galutiniam erdvinį duomenų rinkinyje turi būti identifiukuoti pakitę erdviniai duomenys pradiniam erdvinį duomenų rinkinyje arba siekiant sumažinti harmonizavimo kaštus, kai transformuojami tik pakitę erdviniai duomenys. Šis procesas skirtas tam, kad būtų galima užtikrinti nepertraukiamą erdvinį duomenų teikimą.
4. Erdvinių duomenų transformavimo procesui, kurio metu gauti erdviniai duomenys perkeliama į galutinių erdvinį duomenų bazę ir, kuris priklausomai nuo analizės etape gautų rezultatų, gali apimti:
 - 4.1. Atranką (vienas iš erdvinį duomenų generalizavimo uždavinių):
 - 4.1.1. erdvinį duomenų elementų atranka – atrenkami tik tie elementai iš atrinkto erdvinį duomenų rinkinio, kurie turės būti transformuojami;
 - 4.1.2. elementų savybių atranka – atrenkamos tik tos atrinktų elementų savybės, kurios atitinka galutinį erdvinį duomenų modelį.
 - 4.2. Pervadinimą: erdvinį duomenų elementų ir (arba) jų savybių reikšmių pervadinimas.

- 4.3. Perklasifikavimą: erdvinių duomenų elementų ir (ar) jų savybių perklasifikavimas, taikant galutiniame erdvinių duomenų modelyje naudojamus klasifikatorius.
- 4.4. Apjungimą arba išskaidymą: erdvinių duomenų elementų ir (ar) jų savybių apjungimas arba išskaidymas.
- 4.5. Erdvinių duomenų elementų generalizavimą: erdvinių duomenų elementų kiekio mažinimą (atranką), eliminavimą, konfigūracijos suprastinimą, erdvinių duomenų taškinių, linijinių ar plotinių elementų keitimą iš vieno tipo į kitą (pavyzdžiui, plotinis elementas keičiamas tašku arba jį ribojančia linija); erdvinių duomenų elementų apjungimu ir pan.
- 4.6. Perskirstymą:
 - 4.6.1. erdvinių duomenų rinkinys išskaidomas pagal atskirus erdvinių duomenų elementus, arba elementai pagal atskiras jų savybes;
 - 4.6.2. erdvinių duomenų rinkinys sudaromas apjungiant atskirus erdvinių duomenų elementus arba elementai apjungiami pagal atskiras jų savybes.
- 4.7. Papildymą:
 - 4.7.1. erdvinių duomenų elementų savybių papildymas trūkstamomis reikšmėmis, atsižvelgiant į kitas elemento savybes;
 - 4.7.2. erdvinių duomenų elementų numatytųjų savybių galutiniame erdvinių duomenų rinkinyje priskyrimas.
5. Transformuotų erdvinių duomenų perkėlimui į galutinį erdvinių duomenų rinkinį.
6. Galutinių erdvinių duomenų patikrai. Šis siūlomas papildomas procesas skirtas įsitikinti, ar transformavimas buvo atliktas teisingai. Rastos klaidos turi būti išsaugomos klaidų duomenų bazėje. Išanalizavus šias klaidas būtų galima spręsti, ar transformavimas įvyko neteisingai dėl netikslumų pradiniam erdvinių duomenų rinkinyje, ar dėl neteisingai realizuoto transformavimo proceso.

Erdvinių duomenų transformavimo metu reikia numatyti gautų erdvinių duomenų koordinačių sistemos ar duomenų formato keitimą tik tuo atveju, jei jos skiriasi pradinių ir galutinių erdvinių duomenų rinkiniuose.

Erdvinių duomenų transformavimo procese erdvinių duomenų generalizavimas turi būti atliekamas tais atvejais, kai:

1. to reikalauja semantinis erdvinių duomenų suderinamumas (pavyzdžiui, pradinių erdvinių duomenų rinkinyje yra kaupiami erdviniai duomenys apie sodus, ariamą žemę ir kultūrines ganyklas, tuo tarpu galutiniame erdvinių duomenų rinkinyje erdvinių duomenų elementai identifikuoja dirbamą žemę, neskaidant jos į potipius) arba galutinių erdvinių duomenų elementų kaupimo taisyklės. Pavyzdžiui GDR10LT erdvinių

duomenų elementai vaizduojantys kelių ir gatvių ašinės linijas yra suskaidyti atskirais elementais ne tik priskirtoms skirtingoms savybių reikšmėms, bet ir kai susikerta su kitais elementais, tuo tarpu sudarant kelių tinklo modelį, tokie papildomi sukirtimai gali būti nedaromi. Tokiu atveju transformuojant duomenis turi būti naudojamas elementų generalizavimas, taikant kaimyninių to paties tipo elementų apjungimo algoritmus:

- 1.1 plotiniams elementams, kai jie turi bendrą kraštinę ar yra nutolę vienas nuo kito nustatytu minimaliu atstumu;
- 1.2 linijiniams elementams, kai jų pradžios ar pabaigos taškas sutampa su kito elemento pradžios ar pabaigos tašku, arba elementai yra lygiagretūs vienas kitam ir yra nutolę vieni nuo kitų tam tikru nustatytu minimaliu atstumu, jei tai leidžia galutinių erdvinių duomenų modelis.
- 1.3 Taškiniams elementams, kai jų sanakaupa turi apibūdinti vienas taškinis elementas identifikuojantis sanakupos vietą.
2. Pradiniam ir galutiniam erdvinių duomenų modeliui tapatūs erdviniai objektai vaizduojami skirtingais erdvinių duomenų elementais (2.1 lentelė) arba duomenys yra kaupiami skirtingu detalumu. Pavyzdžiui, GDR10LT yra kaupiamos geležinkelių bėgių kelių ašinės linijos (GDR10LT... 2014), o GDR50LT (GDR50LT... 2012) ar INSPIRE transporto tinkle (INSPIRE data specification on Transport... 2014) – geležinkelio kelio ašinė linija.
3. Ženkliai skiriasi pradinių ir galutinių erdvinių duomenų kartografinis mastelis.

2.1 lentelė. Pagrindinių georeferencinių erdvinių duomenų elementų vaizdavimui naudojami geometrijos tipai

Table 2.1. The geometry types used for representing elements of main (geo) referenced spatial data

Georeferenciniai objektai	Vaizduojamų elementų tipas		
	linijinis	plotinis	taškinis
Keliai, gatvės	menama objekto ašinė linija arba kelio, gatvės kraštai	objekto dengiama teritorija	–
Geležinkeliai	menama objekto ašinė linija	objekto dengiama teritorija	–
Upės, kanalai, melioracijos grioviai	menama objekto ašinė linija arba krantas	objekto dengiama teritorija	–

2.1 lentelės pabaiga

Georeferenciniai objektai	Vaizduojamų elementų tipas		
	linijinis	plotinis	taškinis
Ežerai, tvenkiniai, kūdros	krantas	objekto dengiama teritorija	–
Pastatai	objekto dengiamos teritorijos riba	objekto dengiama teritorija	objekto centro koordinatės
Žemėnaudos	objekto dengiamos teritorijos riba	objekto dengiama teritorija	–
Užstatytos teritorijos	objekto dengiamos teritorijos riba	objekto dengiama teritorija	–
Miškų kvartalinės linijos	menama objekto ašinė linija, objekto ribos linijos	–	–
Geodezinio pagrindo punktai	–	–	objekto centro koordinatės

Kūrimo etape remiantis projektavimo rezultatais turi būti sukurta harmonizavimo technologija, naudojant pasirinktą programinę įrangą ir atsižvelgiant į turimas informacines technologijas, atlikti jos testavimai, diegimas ir dokumentavimas vadovaujantis tipinėmis informacinių technologijų kūrimo metodikomis. Įgyvendinant erdvinės informacijos transformavimo technologiją, kai yra atliekamas erdvinių duomenų generalizavimas, kuriant rekomenduojama lygiagrečiai vykdyti testavimus, pradedant testuoti naudojant nedidelės apimties erdvinis duomenis, testavimo teritoriją vis didinant. Kadangi atliekant erdvinių duomenų generalizavimo modelių, aprašytų 2.2 poskyryje, projektavimą ir realizavimą, nustatyta, kad esant didelės apimties erdvinių duomenų rinkiniams, yra sudėtinga identifikuoti analizės etape visų pradinių erdvinių duomenų elementų specifiškumą, o iškart pradėti sukurtų modelių testavimus naudojant pilnos apimties erdvinis duomenis kartais būna nepraktiška, nes atlikti visas modelyje aprašytas procedūras gali reikėti daug laiko ir technologinių resursų.

2.2. Erdvinių duomenų generalizavimo algoritmai

2.1 poskyryje aptarta, kada erdvinių duomenų harmonizavime turi būti taikomi erdvinių duomenų generalizavimo metodai. Todėl šiame poskyryje analizuojami esami pagrindiniai generalizavimo algoritmai, sprendžiantys pavienius ar grupę

generalizavimo uždavinių. Tačiau pavienių erdvinių duomenų generalizavimo algoritmų naudojimas paprastai neleidžia pasiekti norimo rezultato generalizuojant grupę tarpusavyje susijusių erdvinių duomenų. Ypač tai aktualu generalizuojant erdvinis duomenis vaizduojančius gamtinius ir antropogeninius objektus, kadangi jie dėl savo specifiškumo generalizavimo procese turi išlaikyti tam tikras tik jiems būdingas savybes ir ryšius su kitais elementais. Disertacijos 1.2.3 poskyryje pateikta mokslinės literatūros analizė parodė, kad nėra sukurtų universalių modelių, apimančių skirtingus generalizavimo algoritmus ir erdvinių duomenų modeliavimo funkcijas, kadangi jie priklauso nuo pradinio ir galutinio erdvinių duomenų modelių, erdviniam duomenim keliamų reikalavimų, mastelio ir pan. Šiame poskyryje siūlomi algoritmai, skirti specifiniams generalizavimo uždaviniams spręsti, kuriuos kombinuojant su kitais generalizavimo algoritmais ir erdvinių duomenų modeliavimo funkcijomis, galima pasiekti norimų rezultatų kuriant generalizavimo modelius.

Analizuojant erdvinių duomenų elementų kokybės reikalavimus paprastai yra išskiriami keturi kokybę charakterizuojantys pagrindiniai komponentai (Müller 1991):

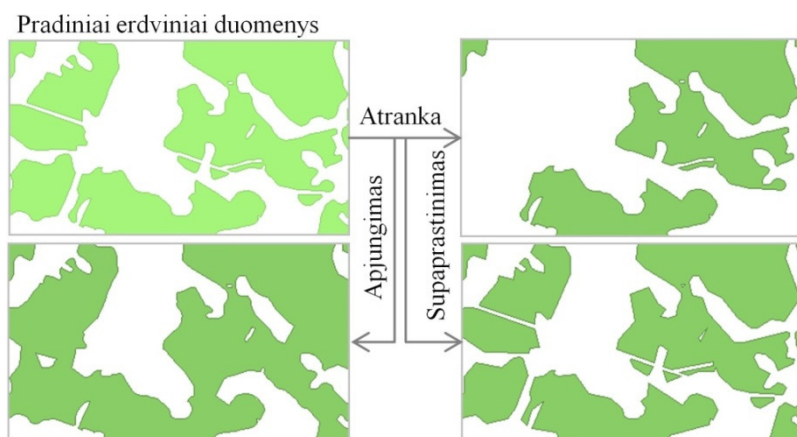
- elemento vietos tikslumas,
- elementams priskirtų savybės reikšmių tikslumas,
- erdvinių duomenų vientisumas,
- erdvinių duomenų pilnumas.

Atliekant erdvinių duomenų generalizavimą kai yra harmonizuojami skirtingo mastelio erdviniai duomenų rinkiniai, dažnai sumažėja erdvinių duomenų elementų padėties tikslumas. Taip pat šis procesas gali įtakoti erdvinių duomenų pilnumą ir vientisumą, kadangi, atlikus informacijos apibendrinimą, gali būti prarandama tam tikra svarbi informacija arba erdvinių duomenų elementai gali būti neteisingai generalizuoti. Todėl generalizavimo procese yra būtina parinkti tinkamus algoritmus ir jų sekas, norint gauti kokybės reikalavimus tenkinančius rezultatus.

Erdvinių duomenų generalizavime yra naudojami įvairūs algoritmai. Mokslinėje literatūroje yra išskiriami trys pagrindiniai faktoriai algoritmams parinkti: situacijos analizės rezultatai, erdvinio objekto tipas (pavyzdžiui, kelias) ir žemėlapių ar erdvinių duomenų rinkinio kartografinis mastelis (Cecconi 2003). Bendruoju atveju disertacijoje išskiriami trys pagrindiniai generalizavimo metodai, kurie įtakoja erdvinių duomenų vaizdo suprastinimą, tai atranka, konfigūracijos suprastinimas ir agregavimas (Papšienė *et al.* 2013b) (2.3 pav.).

Naudojant atranką pagal apibrėžtas kokybines ir kiekybines elemento savybes, sumažinamas bendras erdvinių duomenų kiekis (pavyzdžiui, harmonizavimo metu atrenkami elementai, kurie identifikuoja tik miškų objektus arba upes ilgesnes, nei 10 km, arba „0“ klasės geodezinius punktus). Erdvinių duomenų elementų konfigūracijos suprastinimas vykdomas,

atsižvelgiant į nustatytą maksimalų leidžiamą nuokrypį; o kaimyninių erdvinio duomenų elementų, kuriems priskirtos vienodos savybių reikšmės, apjungimas (agregavimas) vykdomas tuomet, kai tarp kaimyninių elementų yra atstumas mažesnis nei nustatytas minimumas. Taikant šiuos pagrindinius metodus yra eliminuojami erdvinio duomenų rinkiniui ar žemėlapiui nebūdingi ar nereikšmingi erdviniai duomenys ar jų elementai: sumažinamas erdvinio duomenų kiekis, panaikinami per maži atstumai tarp elementų ir suprastinama erdvinio duomenų elementų konfiguracija. Tačiau konkrečių algoritmų praktinis parinkimas turi priklausyti nuo analizės etape gautų rezultatų ir turimos programinės įrangos.



2.3 pav. Plotinių elementų modifikacijos taikant skirtingus generalizavimo metodus
Fig. 2.3. Modifications of polygon features using different methods of generalisation

Tolimesniuose poskyriuose yra aptariami kai kurie algoritmai, taikytini įgyvendinant pagrindinius aukščiau minėtus generalizavimo metodus, taip pat pateikiamas algoritmų realizavimas naudojant ESRI programinės įrangos *ArcGIS* grafinės modeliavimo aplinkos *ModelBuilder* galimybes, nurodant siūlomas naudoti pagrindines *ArcGIS* funkcijas.

2.2.1. Elementų atranka

Erdvinių duomenų atranka paprastai yra pagrįsta ne algoritmu, o tam tikrų taisyklių taikymu (Urbanas 2001a), kurios nurodo kokie erdviniai duomenys turi būti perkeltami į smulkesnio mastelio erdvinio duomenų rinkinį. Kaip aprašyta 2.1 poskyryje pateiktoje patobulintoje erdvinio duomenų harmonizavimo metodikoje, transformavimo procesas paprastai apima pirminę ir antrinę erdvinio duomenų atranką.

Pirminė erdvinių duomenų atranka vykdoma įvertinant erdvinių duomenų elementams priskiriamas savybes, atsižvelgiant į tai, kokie yra pradinių ir galutinių erdvinių duomenų modeliai ir klasifikatoriai, skirti elementų savybėms apibūdinti.

Prieš atliekant pirminę atranką turi būti identifikuojamos erdvinių duomenų elementų klasės, iš kurių bus atrenkami elementai, ir elementų savybių reikšmės pagal kuriuos bus atrenkami reikiami elementai. Tuomet atrankos veiksmų seka, kurioje atsižvelgiama į priskirtas elementams savybių reikšmes, apims veiksmus pateiktus 2.2 lentelėje.

2.2 lentelė. Erdvinių duomenų elementų atrankos, naudojant savybių reikšmes, algoritmas

Table 2.2. The algorithm for selection of spatial data features by using values of attributes

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Įkeliami erdvinių duomenų elementų klasė <i>EK1</i> , kurios elementai turi būti atrinkti.	–
Pagrindinė seka:		
2.	Suformuojama užklausa, kuri atranka elementus iš <i>EK1</i> , turinčius identifikuotas savybių reikšmes.	<i>Select by Attribute.</i>
3.	Atrinkti <i>EK1</i> elementai perkeliama į naują erdvinių duomenų elementų klasę <i>EK2</i> .	<i>Feature Class to Feature Class</i>
4.	Rezultatas: elementų klasė <i>EK2</i> , kurioje yra tik elementai atitinkantys identifikuotas savybių reikšmes.	–
Alternatyvi seka:		
2.	Suformuojama užklausa, kuri atranka elementus, neturinčius identifikuotų savybių reikšmių.	<i>Make Feature Layer, Select by Attribute.</i>
3.	Panaikinami atrinkti elementai.	<i>Delete Features</i>
4.	Rezultatas: elementų klasė <i>EK1</i> , kurioje yra tik elementai atitinkantys identifikuotas savybių reikšmes.	–

Antrinė erdvinių duomenų atranka turi būti vykdoma atliekant erdvinių duomenų harmonizavimą tarp skirtingo kartografinio mastelio rinkinių. Kadangi smulkėjant kartografiniui masteliui tiek žemėlapyje, tiek erdvinių duomenų rinkinyje yra saugomas ir vaizduojamas mažesnis kiekis elementų. Todėl atrankos metu papildomai turi būti įvertinamas vaizduojamų erdvinių duomenų elementų kiekis ir (ar) tarpusavio elementų padėtis.

Optimalaus kiekio nustatymas yra sudėtingas uždavinys, priklausantis ne tik nuo žemėlapių ar erdvinio duomenų rinkinio kartografinio mastelio, bet ir nuo erdvinio duomenų rinkinio ar žemėlapių paskirties ar vartotojų poreikių.

Klasikinėje kartografijoje kuriant žemėlapius, kurių erdviniams duomenims paruošti naudojami stambesnio mastelio erdviniai duomenys, elementų atrankai taikytinas pagrindinis erdvinio duomenų elementų atrankos principas, kurį pasiūlė Töpfer ir Pillewizer (1966). Šis principas leidžia nustatyti ryšį tarp žemėlapių mastelio ir elementų kiekio žemėlapyje, tai yra parodo kiek elementų reikėtų perkelti iš stambesnio mastelio erdvinio duomenų rinkinio į smulkesnį mastelį. Šis ryšys išreiškiamas formule:

$$n_c = n_s \sqrt{\frac{S_c}{S_s}}, \quad (2.1)$$

čia n_c yra elementų skaičius žemėlapyje, kurio mastelio vardiklis S_c , o n_s – elementų skaičius pradiniam stambesnio mastelio žemėlapyje, kuris yra naudojamas pirmojo atnaujinime ir, kurio mastelio vardiklis yra S_s .

Tačiau formulės (2.1) negalima vienareikšmiškai taikyti visais atvejais atliekant erdvinio duomenų kartografinį generalizavimą, t. y. kuriant žemėlapi, kadangi:

- skirtinguose žemėlapiuose erdvinio duomenų elementai gali būti vaizduojami naudojant įvairaus dydžio, pločio sutartinius ženklus;
- erdvinio duomenų elementų kiekis, kaip paminėta anksčiau, žemėlapyje gali priklausyti nuo žemėlapių paskirties ir (ar) vartotojų poreikių.

Todėl Töpfer formulę (2.1) praplėtė įvesdamas dvi konstantas, atsižvelgiant į žemėlapių paskirtį ir naudojamus sutartinius ženklus erdviniams duomenims atvaizduoti:

$$n_c = n_s C_e C_f \sqrt{\frac{S_c}{S_s}}, \quad (2.2)$$

čia C_e yra žemėlapių elemento simbolio dydžio konstanta, C_f – žemėlapių elemento simbolio formos konstanta.

Konstantos C_e reikšmė, priklauso nuo abstrakčiai apibrėžiamo žemėlapyje naudotino sutartinio ženklo dydžio:

- elementams, kuriems pavaizduoti žemėlapyje naudojamas sutartinis ženklas, nėra išdidinamas, t. y. vaizduojami erdvinio duomenų elementai matomi žemėlapyje jų neišdidinus, koeficientas $C_{e1} = 1,0$;

- plotiniams elementams, kurių riba vaizduojama linijiniu simboliu, kuris nėra sąlyginai išdidinamas (pavyzdžiui, ežerams), koeficiento reikšmė skaičiuojama pagal formulę:

$$C_{e2} = \sqrt{\frac{S_s}{S_c}}; \quad (2.3)$$

- elementams, kuriems pavaizduoti žemėlapyje naudojamas išdidintas simbolis, koeficientas skaičiuojamas pagal formulę:

$$C_{e3} = \sqrt{\frac{S_c}{S_s}}. \quad (2.4)$$

Konstantos C_f reikšmė, priklauso nuo žemėlapyje naudojamo sutartinio ženklo formos:

- elementams, kuriems pavaizduoti žemėlapyje naudojamas sutartinis ženklas nėra išdidinamas, tai yra vaizduojami erdvinių duomenų elementai matomi žemėlapyje jų neišdidinus, koeficientas $C_{f1} = 1,0$;
- linijinių elementų sutartiniam ženklu, kai generalizavimo procese būtina atsižvelgti į linijinio simbolio plotį, koeficientas skaičiuojamas pagal formulę:

$$C_{f2} = \left(\frac{w_s}{w_c}\right) \sqrt{\frac{S_s}{S_c}}, \quad (2.5)$$

čia w_s – linijinio elemento plotis pradiniam žemėlapyje, w_c – linijinio elemento plotis naujai kuriamame žemėlapyje;

- plotinių elementų sutartiniam ženklu, kai generalizavimo procese būtina atsižvelgti į plotinio simbolio užimamą plotą, koeficientas skaičiuojamas pagal formulę:

$$C_{f3} = \left(\frac{a_s}{a_c}\right) \sqrt{\frac{S_s}{S_c}}, \quad (2.6)$$

čia a_s – plotinio elemento plotas pradiniam žemėlapyje, a_c – plotinio elemento plotas naujai kuriamame žemėlapyje.

Töpfer siūlomas elementų kiekio įvertinimo principas leidžia kontroliuoti atrenkamų erdvinį duomenų kiekį, tačiau negali atsakyti, kurie konkrečiai elementai turi būti atrenkami, be to atrankos principas yra tinkamas tik tuo atveju, kai stambesnio mastelio erdvinį duomenų rinkinys atitinka erdvinį duomenų pilnumo kokybės komponentą.

Šios atrankos principų taikymo galimybės įvertintos analizuojant:

- 2012 metų GDR10LT;
- 2012 metų GDR50LT, kuris sukurtas ir atnaujintas rankiniu būdu naudojant GDR10LT;
- GDR50LT, kuris buvo automatizuotai sukurtas 2013 metais vykdant Nacionalinės žemės tarnybos prie Žemės ūkio ministerijos užsakytą GDR50LT atnaujinimo galimybių tyrimą, naudojant automatizuotus generalizavimo metodus (Papšienė *et al.* 2013a) (tyrimo vykdytojas VĮ Distancinių tyrimų ir geoinformatikos centras „GIS-Centras“. Šį tyrimą autorė atliko kartu su kolegomis).

Analizės metu nustatytas erdvinį duomenų elementų kiekis, esantis tam tikroje apibrėžtoje teritorijoje GDR50LT, kuris buvo sukurtas ir atnaujintas rankiniu būdu, ir kuris buvo sukurtas automatizuotai, taikant generalizavimo metodus. Nustatytas elementų kiekis palygintas su teoriškai išskaičiuotu optimaliu elementų kiekiu, atsižvelgiant į GDR10LT esantį elementų kiekį toje pačioje teritorijoje. Įvertinimo rezultatai pateikti 2.3 lentelėje.

Šiuo atveju formulėje (2.2) nurodytas koeficientas $C_e = C_{el}$, o $C_f = C_{fl}$, kadangi erdviniam duomenų rinkinyje elementai yra kaupiami neatsižvelgiant į galimus elementų vaizdavimo žemėlapyje ypatumus, t. y. į konkrečiame žemėlapyje naudojamų sutartinių ženklų dydį. Atsižvelgiant į gautus rezultatus matome, kad GDR50LT erdvinį duomenų elementų kiekis testuojamoje teritorijoje yra paprastai didesnis, nei rekomenduojamas, tačiau ne radikaliai, išskyrus hidrografijos, geležinkelių ir miškų elementų kiekį, gautą atnaujinus GDR50LT rankiniu būdu. Tai reiškia, kad galutinį elementų kiekį įtakoja erdvinį duomenų vartotojų poreikis turėti detalesnius duomenis, nei įprastiniuose žemėlapiuose. Tačiau atsižvelgiant į analizės rezultatus siūloma formulę (2.3) taikyti preliminariai įvertinant, koks erdvinį duomenų tankumas galėtų būti vaizduojamas atlikus skirtingo mastelio erdvinį duomenų rinkinių harmonizavimą, kai galutiniame erdvinį duomenų modelyje nėra aiškiai apibrėžtas pageidaujamas erdvinį duomenų elementų tankumas.

Erdvinių duomenų tankumo mažinimas priklauso nuo to, kokius erdvinis objektus jie vaizduoja, o taip pat nuo galutiniams erdviniams duomenims taikomų reikalavimų.

2.3 lentelė. Erdvinių duomenų elementų kiekio įvertinimas**Table 2.3.** Evaluation of number of spatial data features

Erdvinių duomenų elementų klasė	Elementų skaičius GDR10LT	Teorinis elementų skaičius 1:50 000	Elementų skaičius po automatinio generalizavimo GDR50LT	Elementų kiekio skirtumas tarp teorinio elementų skaičiaus ir generalizuotų automatinio būdu, %	Elementų skaičius po rankinio generalizavimo GDR50LT	Elementų kiekio skirtumas tarp teorinio elementų skaičiaus ir generalizuotų rankiniu būdu, %
1.	2.	3.	4.	8.	5.	9.
Hidrografija	18481	8316	10069	21,08%	13939	67,62%
Keliai	28215	12697	15834	24,71%	17106	34,72%
Užstatytos teritorijos	14457	6506	5584	14,17%	–	0,00%
Geležinkeliai	590	266	236	11,28%	133	50,00%
Ežerai, tvenkiniai	7499	3375	2894	14,25%	2138	36,65%
Mišakai	6847	3592	2831	21,19%	1498	58,30%

Bendruoju atveju, jei erdvinių duomenų specifikacijose nėra numatyta kitaip, projektuojant transformavimo procesą, kuriame turi būti mažinamas erdvinių duomenų kiekis, priklausomai nuo elementų tarpusavio išsidėstymo ir sąsajos su elementais kitose elementų klasėse rekomenduojama:

1. Nustatyti atrenkamų elementų, esančių toje pačioje erdvinių duomenų elementų klasėje prioritetus, skirtus elementams atrinkti. Nurodyti, kurie elementai negali būti eliminuojami ir, kurie, atsižvelgiant į jų svarbumą, gali būti eliminuojami.
2. Nurodyti, kurie elementai kitose elementų klasėse gali įtakoti erdvinių duomenų elementų eliminavimą arba atrinkimą.
3. Nustatyti mažiausią leidžiamą atstumą tarp atrenkamų elementų.

Atrenkamų erdvinių duomenų elementų prioritetai turi būti nustatomi atsižvelgiant į erdvinių duomenų rinkinio paskirtį. Prioritetai sudaromi atsižvelgiant į elementams priskirtas savybių reikšmes (pavyzdžiui, keliams prioritetai gali būti išdėstomi atsižvelgiant į jiems priskiriamą valstybinės ir vietinės reikšmės kelių kategorijas), elementų plotą arba ilgį. Taip pat papildomai gali būti įvertinama elementų tarpusavio padėtis (pavyzdžiui, geodeziniai punktai prioritetas gali būti išdėstomas atsižvelgiant į jų klasę ir

punkto tipą, papildomai įvertinant atstumą iki kelių sankryžos, t. y. esantys arčiau sankryžos gali turėti didesnę prioritetą, nes jie yra lengviau pasiekiami, nei esantys tos pačios klasės toliau nuo sankryžos).

Elementų tarpusavio padėtį reikėtų vertinti:

1. atstumu tarp elementų toje pačioje ar skirtingose elementų klasėse;
2. kryptimi, kuria vienas elementas yra nutolęs kito elemento atžvilgiu;
3. kampu tarp dviejų linijinių elementų toje pačioje erdvinių duomenų elementų klasėje;
4. topologiniais ryšiais tarp elementų (pavyzdžiui, upė jungianti du ežerus).

Visi erdvinių duomenų elementai susideda iš x , y koordinatinių poros ar jų sekos. Paprastai norint įvertinti atstumą tarp dviejų elementų, nepriklausomai nuo jo geometrijos, yra vertinamas atstumas tarp dviejų skirtingų elementų taškų. Atstumas S tarp taškinių elementų, kurių koordinatės x_1 , y_1 ir x_2 , y_2 , yra apskaičiuojamas naudojant formulę:

$$S = \sqrt{(\Delta x)^2 + (\Delta y)^2}, \quad (2.7)$$

čia Δx ir Δy koordinatinių prieaugiai $\Delta x = x_2 - x_1$ ir $\Delta y = y_2 - y_1$.

Norint atrinkti elementus, kurie vienas nuo kito yra nutolę nustatytu mažesniu, nei leistinas atstumas, taikytinas algoritmas pateiktas 2.4 lentelėje.

2.4 lentelė. Erdvinių duomenų elementų atrankos, atsižvelgiant į atstumus tarp elementų, algoritmas

Table 2.4. The algorithm for selection of spatial data features taking in to account the distance between features

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Įkeliama erdvinių duomenų elementų klasė <i>EK1</i> , kurios elementai turi būti atrinkti.	–
2.	Įkeliama erdvinių duomenų elementų klasė <i>EK2</i> , pagal kurios elementus turi būti atrinkti <i>EK1</i> elementai (atliekama, jei reikia atrinkti <i>EK1</i> elementus priklausomai nuo atstumo iki <i>EK2</i> elementų).	–
Pagrindinė seka:		
3.	Nustatytu paieškos spinduliu, kurio ilgis lygus mažiausiam leistinam atstumui tarp elementų, randami kaimyniniai elementai.	<i>Generate Near Table</i>

2.4 lentelės pabaiga

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
	<p>Jei reikia nustatyti atstumą tarp elementų esančių:</p> <ul style="list-style-type: none"> – toje pačioje elementų klasėje, tuomet atstumai nustatomi tarp <i>EK1</i> elementų; – tarp skirtingų elementų klasių, tuomet atstumai nustatomi tarp <i>EK1</i> ir <i>EK2</i> elementų. <p>Paskaičiuojamas atstumas tarp rastų kaimyniniu elementų taikant (2.7) formulę.</p> <p>Rezultatas: lentelė, kurioje nurodytos elementų poros ir tarp jų nustatyti atstumai.</p>	
4.	<i>EK1</i> apjungiama su lentele, gauta trečiame žingsnyje, naudojant identifikatorius, kurie unikaliai identifikuoja elementus <i>EK1</i> ir minėtoje lentelėje.	<i>Add Join</i>
5.	Analizuojamos poros elementų, tarp kurių atstumas yra mažesnis, nei leidžiamas nustatytas minimumas. Elementas, turintis mažesnę prioritetą, nei kitas poros elementas, panaikinamas.	<i>Select by Attribute</i> <i>Delete Feature</i>
Alternatyvi seka:		
3.	Nustatytu atstumu <i>S</i> , kurio ilgis lygus pusei mažiausiai leistinam atstumui tarp elementų, sukuriama buferinės zonos aplink <i>EK1</i> elementus.	<i>Buffer</i>
4.	Nustatytu atstumu <i>S</i> , kurio ilgis lygus pusei mažiausiai leistinam atstumui tarp elementų, sukuriama buferinės zonos aplink <i>EK2</i> elementus (atliekama, jei reikia atrinkti <i>EK1</i> elementus priklausomai nuo atstumo iki <i>EK2</i> elementų).	<i>Buffer</i>
5.	Atrenkamos sukurtos buferinės zonos, kurios kertasi viena su kita. Šiam veiksmui atlikti siūloma buferines zonas apjungti. Kertančios viena kitą zonos, apjungus sudarys didesnio ploto teritoriją, nei sukurtų trečiame žingsnyje.	<i>Union, Dissolve, Select by Attribute.</i>
6.	Atrenkami elementai, kurie patenka į atrinktas penktame žingsnyje zonas.	<i>Select by Location</i>
7.	Analizuojami atrinkti elementai. Elementai, turintys mažesnę prioritetą, panaikinami.	<i>Select by Attribute</i> <i>Delete Feature</i>
	<p>Rezultatas: elementų klasė <i>EK1</i>, kurioje yra elementai:</p> <ul style="list-style-type: none"> – tarp kurių atstumas yra didesnis, nei nustatytas leidžiamas minimumas. 	–

Abejose veiksmų sekose numatyta panaikinti tą elementą, kurio prioriteto reikšmė yra mažesnė. Tačiau tuo atveju, jei atrinktų elementų prioriteto reikšmės

yra vienodos, tuomet siūloma panaikinti atsitiktinai atrinktą elementą. Tuo atveju, kai prioriteto reikšmė parodo, kad nei vienas elementas negali būti panaikintas, tuomet siūloma palikti visus atrinktus elementus, nors tarp jų atstumas ir yra mažesnis nei leistinas. Šiuo atveju, kuriant kartografinę produkciją turėtų būti naudojama generalizavimo operacija, perstumianti elementą iš originalios vietos į kitą, tiek, kad atstumas tarp elementų būtų ne mažesnis, nei leidžiamas minimumas.

2.4 lentelėje pateiktoje pirmoje sekoje yra aprašyti veiksmai leidžiantys analizuoti atstumus tik tarp dviejų elementų, o antroje sekoje analizuojama elementų grupė, tarp kurių kaimyninių elementų atstumas yra mažesnis, nei nustatytas minimumas.

Tačiau pastarojoje sekoje analizuojamoje grupėje gali būti elementų tarp kurių atstumas bus didesnis, nei nustatytas leidžiamas minimumas, t. y. jie bus nutolę vienas nuo kito leidžiamu atstumu, jei elementai išsidėstę išilgai vienas kito.

Norint nustatyti kryptį, kuria vienas elementas yra nutolęs kito elemento atžvilgiu, galima įvertinti apskaičiuojant linijos, einančios per analizuojamus taškus, direkcinį kampą α :

$$tg \alpha = \frac{\Delta y}{\Delta x} . \quad (2.8)$$

Norint atrinkti elementus, kurie vienas nuo kito yra nutolę tam tikra kryptimi, taikytinas algoritmas pateiktas 2.5 lentelėje.

2.5 lentelė. Erdvinių duomenų elementų atrankos, atsižvelgiant į kryptis tarp elementų, algoritmas

Table 2.5. The algorithm for selection of spatial data features taking in to account the directions between features

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Įkeliama erdvinių duomenų elementų klasė <i>EK1</i> , kurios elementai turi būti atrinkti.	–
2.	Įkeliama erdvinių duomenų elementų klasė <i>EK2</i> , pagal kurios elementus turi būti atrinkti <i>EK1</i> elementai. Šis žingsnis atliekamas, jei reikia atrinkti <i>EK1</i> elementus, kurie yra nutolę nuo elementų esančių elementų klasėje <i>EK2</i> .	–

2.5 lentelės pabaiga

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
3.	Nustatyti paieškos spinduliu, randami kaimyniniai elementai. Jei reikia nustatyti atstumą tarp elementų esančių: – toje pačioje elementų klasėje, tuomet atstumai nustatomi tarp <i>EK1</i> elementų; – tarp skirtingų elementų klasių, tuomet atstumai nustatomi tarp <i>EK1</i> ir <i>EK2</i> elementų. Nustatoma kryptis tarp rastų kaimyniniu elementų taikant (2.8) formulę. Rezultatas: lentelė, kurioje nurodytos elementų poros ir tarp jų nustatytos kryptys.	<i>Generate Near Table</i>
4.	Trečiame žingsnyje gautoje lentelėje analizuojamos taškų poros ir tarp jų nustatytos kryptys. Atrenkamos taškų poros, kurių kryptys, patenka į pageidaujamų krypčių intervalą. Atrinktos taškų poros perkeliama į naują lentelę.	<i>Select by Attribute</i>
5.	<i>EK1</i> apjungiamas su lentele, gauta trečiame žingsnyje, naudojant identifikatorius, kurie unikaliai identifikuoja elementus <i>EK1</i> ir minėtoje lentelėje.	<i>Add Join</i>
	Rezultatas: elementų klasė <i>EK1</i> , kurioje yra elementai: – tarp kurių atstumas yra didesnis, nei nustatytas leidžiamas minimumas.	–

Tačiau erdvinį duomenų elementų atrankos nereikėtų vykdyti tik įvertinant nustatytą kryptį, papildomai reikėtų vertinti ir atstumą tarp dviejų elementų.

Norint nustatyti kampą tarp dviejų linijinių elementų galima naudoti modifikuotą 2.5 lentelėje aprašytą algoritmą: trečiame žingsnyje yra paliekamos visos elementų poros, ketvirtame žingsnyje nustatomas kampas tarp linijinių elementų atėmus vieną iš kitos direkcinų kampų reikšmes.

Atliekant erdvinį duomenų elementų atranką, atsižvelgiant į elementų topologines sąsajas, turi būti analizuojamą elementų tarpusavio padėtis. Šiuo atveju siūloma analizuoti, ar

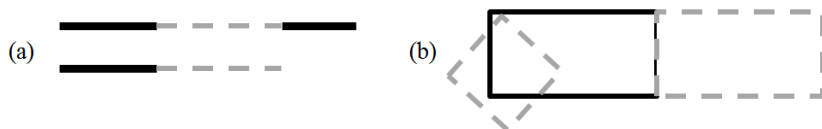
- vienas elementas liečia kitą elementą;
- vienas elementas pilnai ar dalinai dengia kitą elementą;
- vieno elemento kraštinė dengia kito elemento raštinę;
- vienas elementas yra kito elemento viduje.

2.2.2. Elementų eliminavimas ir apjungimas

Erdvinių duomenų elementų eliminavimas yra taikomas erdvinių duomenų generalizavime, kai elementai, netenkinantys jiems keliamų reikalavimų yra pašalinami. Tačiau šiuo atveju gali būti neišlaikytas elementų vientisumas erdvinių duomenų elementų klasėje. Pavyzdžiui, panaikinus linijinį elementą, kurio pradžia ir pabaiga sutampa su kaimyninių elementų pradžia ar pabaiga, arba plotinį elementą, kurio visos kraštinės yra bendros su kaimyniniais plotiniais elementais, gali likti tarpas tarp elementų. Todėl pasirenkant tinkamą eliminavimo būdą būtina atsižvelgti, ar jį eliminavus, nebus pažeistas elementų vientisumas ir kokybė erdvinių duomenų elementų klasėje.

Eliminuojamų elementų apjungimo su kaimyniniais elementais algoritmas gali priklausyti, nuo elemento tipo ir jo sąsajos su kaimyniniais elementais toje pačioje erdvinių duomenų elementų klasėje. Išanalizavus galimus eliminavimo būdus, nustatyta, kad eliminavimo procese susiję elementai yra tokie, kai:

1. vieno linijinio elemento pradžia ar pabaiga sutaps su kito linijinio elemento pradžia ar pabaiga (2.4 a);
2. plotinio elemento kraštinės daugiau, kaip vienos viršūnės koordinatės sutaps su kito plotinio elemento viršūnių koordinatėmis (2.4 b).



2.4 pav. Susijusių linijinių (a) ir plotinių (b) elementų pavyzdžiai

Fig. 2.4. Examples of related features of lines (a) and polygons

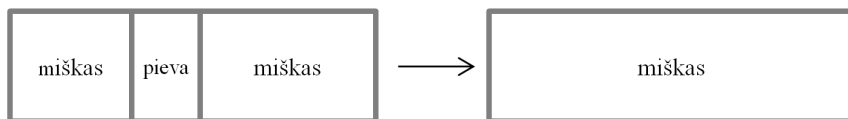
Šiuo atveju taškiniai elementai, kurių koordinatės yra vienodos, nebus susiję. Jie bus identiški ir tolimesnės įtakos į taškinio elemento eliminavimo būdai neturės.

Eliminuojant perteklinius elementus, rekomenduojama juos panaikinti tik tuo atveju, jei jų viršūnių, pradžios ir pabaigos taškų x, y koordinatės nesutampa su kito elemento toje pačioje klasėje viršūnių, pradžios ar pabaigos taškų x, y koordinatėmis. Tais atvejais, kai elemento pradžios ar pabaigos taško x, y koordinatės sutampa su kito elemento pradžios ar pabaigos taško koordinatėmis, priklausomai nuo poreikio ir sprendžiamo uždavinio, perteklinis elementas gali būti ištrintas arba apjungtas su kaimyniniu elementu. Pavyzdžiui, naikinant kabančias kelių ar hidrografijos elementų atšakas, tokie elementai yra panaikinami. Jei analizuojamo elemento pradžios ir pabaigos taškų x, y koordinatės sutampa su kitų elementų pradžios ir pabaigos koordinatėmis,

tuomet elementas privalo būti prijungtas prie kaimyninio elemento, turinčio didesnę prioritetą, arba padalinus pusiau, prie abiejų kaimyninių elementų.

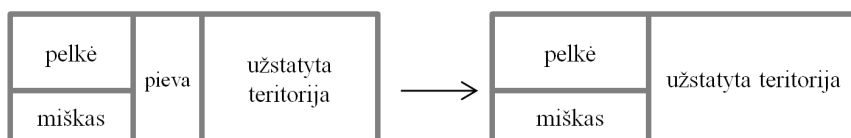
Jei plotinių elementų klasėje tarp elementų gali būti tarpai, tuomet perteklinius elementus rekomenduojama panaikinti, nebent jie yra tarp elementų, kurie nėra pertekliniai. Šiais atvejais ar tais atvejais, kai plotiniai elementai erdvinių duomenų elementų klasėje negali turėti tarpų tarp elementų, eliminavimo atveju elementus būtina apjungti su kaimyniniais plotiniais elementais. Kai eliminuojamo plotinio elemento kraštinė yra bendra su:

1. vienu ar keliais plotiniais elementais, turinčiais tas pačias savybių reikšmes, jis turi būti apjungiamas su gretimais elementais, sudarant vieną plotinį elementą (2.5 pav.);
2. keliais plotiniais elementais, turinčiais skirtingas savybių reikšmes, jis gali būti apjungiamas su vienu iš gretimų elementų arba su keliais gretimais elementais, jo plotą padalinus ir prijungus atitinkamai prie kaimyninių elementų (2.6 pav.).



2.5 pav. Plotinio elemento eliminavimo pavyzdys, kai kaimyniniai elementai turi tokias pačias savybių reikšmes

Fig. 2.5. Example of elimination of polygon where neighbouring elements have the same characteristics



2.6 pav. Plotinio elemento eliminavimo pavyzdys, kai kaimyniniai elementai turi skirtingas savybių reikšmes

Fig. 2.6. Example of elimination of polygon where neighbouring elements have different characteristics

Antruoju aprašytu atveju elementas, turintis daugiau nei vieną kaimyninį elementą, gali būti apjungiamas su gretimu elementu atsižvelgiant į gretimų elementų kokybines ar kiekybines savybes. Jis gali būti apjungiamas su tuo elementu, kurio savybės reikšmė yra svarbesnė arba plotas didesnis, arba ilgesnė bendra kraštinė tarp elementų. Tačiau šiuo atveju rekomenduojama eliminuoti

tik nedidelius plotinius elementus, nes priešingu atveju neproporcingai gali padidėti plotas to elemento, prie kurio buvo prijungtas eliminuojamas elementas. Nors priklausomai nuo vartotojų poreikių antruoju atveju eliminuojamas elementas gali būti prijungiamas ir prie vieno iš kaimyninių elementų, kurio savybės reikšmė yra semantiškai panašesnė į pastarojo, jo neišdalinant skirtingiems elementams.

Erdvinių duomenų agregavimo operacijos gali būti taikomos 2.6 paveiksle pavaizduotai situacijai išspręsti. Jones su kolegomis (1995) pasiūlė spręsti taikant Delaunė trianguliaciją, kurios metu eliminuojama teritorija padalinama į trikampių. Šiuo atveju eliminuojama teritorija padalinama į kelias dalis, kurios prijungiamos prie kaimyninių teritorijų. Atsižvelgiant į galimybę panaudoti esamus erdvinių duomenų rinkinyje linijinius elementus eliminuojamo plotinio elemento padalinimui, šiuo atveju rekomenduoja operuoti linijiniais elementais, ribojančiais visus plotinius elementus, kurie yra pratęsimi iki per eliminuojamą teritoriją einančios egzistuojančio linijinio elemento. Pavyzdžiui, generalizavimo procese yra eliminuojama kelio teritorija, pratęsiant plotinių elementų linijas iki kelio ašinę liniją vaizduojančio linijinio elemento. Šiuo atveju, ne tik eliminuojamas elementas, bet ir jo plotą prijungus prie kaimyninio elemento, pastarojo kraštinė sutaps su egzistuojančiu linijiniu elementu. Algoritmas, skirtas plotinių elementų eliminavimui, juos apjungiant su kaimyniniais elementais, yra aprašytas 2.6 lentelėje. Algoritme turi būti naudojama plotinių elementų erdvinių duomenų elementų klasė, kurioje pertekliniai elementai turi būti eliminuojami, ir plotinius elementus ribojančių linijinių elementų klasė. Šioje elementų klasėje linijiniai elementai vienas su kitu gali liestis tik linijos pradžios ar pabaigos tašku, t. y. linijinis elementas, kurio viršūnės koordinatės sutampa su kito linijinio elemento viršūnės ar pradžios, pabaigos taško koordinatėmis, turi būti padalintas į du atskirus linijinius elementus, taip pat elementai negali dubliuotis. Papildomai algoritme naudojami linijiniai elementai, vaizduojantys eliminuojamų plotinių elementų centrinę liniją. Poskyryje 2.2.5 yra aptarti algoritmai, kurie sukuria plotinių elementų centrinę liniją.

2.6 lentelė. Pavienių linijinių ir plotinių elementų eliminavimo algoritmas

Table 2.6. The algorithm of elimination of single lines or polygons features

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Įkeliami erdvinių duomenų plotinių elementų klasė (<i>EK1_plotiniai</i>), linijinių elementų klasė (<i>EK1_linijiniai</i>).	–
2.	Elementų klasėje <i>EK1_plotiniai</i> atrenkami elementai, kurie turi būti eliminuojami. Naudojamas algoritmai aprašyti 2.2.1 poskyryje.	–

2.6 lentelės tęsinys

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
3.	Atrinktiems elementams priskiriama savybės reikšmė „eliminuoti“.	<i>Add field,</i> <i>Calculate Field.</i>
4.	Pasirenkami visi elementai, esantys elementų klasėje <i>EK1_linijiniai</i> . Iš atrankos pašalinami elementai, kurie turi bendrą kraštinę su plotiniais elementais turinčiais savybės reikšmę „eliminuoti“.	<i>Select by Attribute,</i> <i>Select by Location (BOUNDARY_TOUCHES,</i> <i>remove from selection).</i>
5.	Ketvirtajame žingsnyje atrinkti elementai perkeliami į naują erdvinių duomenų elementų klasę <i>EK2_linijiniai</i> . Papildomai į šią elementų klasę perkeliama <i>EK1_linijiniai</i> elementai. Rezultatas: erdvinių duomenų elementų klasė <i>EK2_linijiniai</i> .	<i>Feature Class to Feature Class,</i> <i>Append</i>
6.	Analizuojama elementų klasė <i>EK2_linijiniai</i> . Sukuriami taškiniai elementai <i>EK2_linijiniai</i> elementų pradžios arba pabaigos taškuose, jei jie neturi bendros pradžios ar pabaigos su kitu linijiniu elementu <i>EK2_linijiniai</i> elementų klasėje. Rezultatas: erdvinių duomenų elementų klasė <i>EK1_taškiniai</i> .	<i>Feature Vertices to Point (dangle)</i>
7.	Atrenkami elementai, esantys <i>EK2_linijiniai</i> , kurie liečia elementus <i>EK1_taškiniai</i> , t. y. atrenkamos linijos, kurių pradžia ir (arba) pabaiga neturi bendros pradžios ir (ar) pabaigos su kaimyniniu elementu. Vykdomas atrinktų linijinių elementų prailginimas tiese iki artimiausio kaimyninio elemento, esančio nustatytu atstumu.	<i>Select by Location, Extend Line</i>
8.	Nustatytu atstumu sutankinamos <i>EK1_linijiniai</i> elementų viršūnės. Iš elemento viršūnių sukuriami taškiniai elementai, kurie išsaugomi elementų klasėje <i>EK2_taškiniai</i> .	<i>Densify, Feature Vertices to Point</i>
9.	Pakartotinai analizuojama elementų klasė <i>EK2_linijiniai</i> . Sukuriami taškiniai elementai <i>EK2_linijiniai</i> elementų pradžios arba pabaigos taškuose, jei jie neturi bendros pradžios ar pabaigos su kitu linijiniu elementu <i>EK2_linijiniai</i> elementų klasėje. Šis pakartotinis procesas yra reikalingas, jei prieš tai atrinktos linijos, pratęsus jas, nesusikirto su centrine linija. Rezultatas: erdvinių duomenų elementų klasė <i>EK3_taškiniai</i> .	<i>Feature Vertices to Point (dangle)</i>
8.	Atliekama artimiausių taškinių elementų paieška tarp elementų esančių elementų klasėse <i>EK2_taškiniai</i> ir <i>EK3_taškiniai</i> . Rezultatas: lentelė, kurioje nurodytos artimiausių elementų poros.	<i>Generate Near Table</i>

2.6 lentelės pabaiga

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
9.	<i>EK2_linijiniai</i> apjungiami su lentele, gauta aštuntame žingsnyje, naudojant identifikatorius, kurie unikaliai identifikuoja <i>EK2_linijiniai</i> elementus ir elementus minėtoje lentelėje. Sukuriamos linijos tarp rastų kaimyninių elementų., kurios išsaugomos elementų klasėje <i>EK3_linijiniai</i> .	<i>Add Join</i> <i>Add XY</i> <i>Calculate</i> <i>Create line from XY</i>
10.	<i>EK3_linijiniai</i> elementai integruojami į elementų klasę <i>EK2_linijiniai</i> .	<i>Append</i>
11.	Elementų klasėje <i>EK1_plotiniai</i> atrenkami elementai, kurie neturi būti eliminuojami. Naudojamas algoritmas aprašytas 2.2 lentelėje.	–
12.	Iš vienuoliktame žingsnyje atrinktų elementų sukuriama taškiniai elementai, vaizduojanti plotinių elementų centroidę, ir išsaugantys plotinių elementų savybių reikšmes. Rezultatas: elementų klasė <i>EK4_taškiniai</i> .	<i>Feature to Point</i>
13.	Iš elementų klasėje <i>EK2_linijiniai</i> elementų sukuriama plotiniai elementai. Rezultatas: elementų klasė <i>EK2_plotiniai</i> .	<i>Feature to Polygon</i>
14.	Tryliktame žingsnyje sukurtiems elementams priskiriamos savybių reikšmės taikytos ankstesniems plotiniams elementams saugomiems elementų klasėje <i>EK1_plotiniai</i> . Šiuo atveju <i>EK2_plotiniai</i> elementams priskiriamos savybių reikšmės <i>EK4_taškiniai</i> elementų, kurie kerta pastaruosius.	<i>Spatial Join</i> <i>Calculate Field</i>
15.	Rezultatas: elementų klasė <i>EK2_plotiniai</i> , kurioje eliminuoti elementai apjungti su kaimyniniais elementais.	–

Algoritmas, aprašytas 2.6 lentelėje, leidžia išlaikyti visame procese bendras elementų kraštines, tačiau gali likti elementų, kuriems nebus priskirtos ankstesnių elementų savybių reikšmės. Jei tokio elemento plotas yra nedidelis, rekomenduojama jį eliminuoti apjungiant su tuo kaimyniniu elementu, su kuriuo jis turi ilgiausią bendrą kraštinę.

2.2.3. Linijinių ir plotinių elementų suprastinimas

Erdvinių duomenų suprastinimas yra dažnai naudojamas erdvinių duomenų atnaujinime, kadangi tiesiogiai įtakoja generalizuojamų duomenų elementų konfigūraciją. Daugelis autorių nagrinėja linijinių, tame tarpe ir plotinių elementų ribų, suprastinimo galimybes ir siūlo savo sprendimus tokių elementų suprastinimui (Berg *et al.* 1998; Cromley *et al.* 1990; Douglas, Peucker 1973; McMaster 1989; Urbanas 2011a, 2011b). Analizuojant esamus pagrindinius

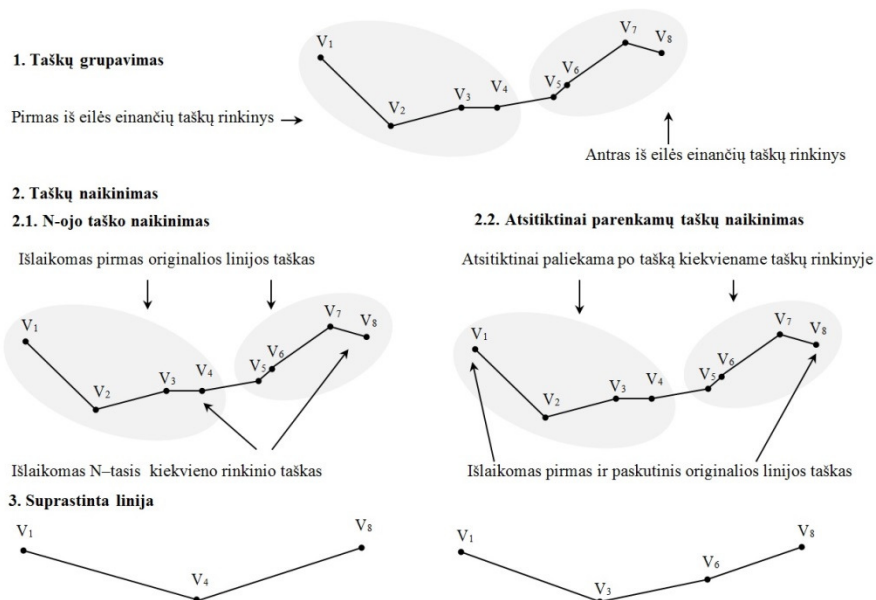
suprastinimo algoritmus, kurie yra įdiegti ir naudojami GIS programinėse įrangose, atsižvelgta į kitų mokslininkų (Shi, Cheung 2006; Zhu *et al.* 2011) atliktus supaprastinimo algoritmų tyrimus.

McMaster (1987) išskiria penkis pagrindinius suprastinimo algoritmų tipus: nepriklausomų taškų algoritmai, vietinio apdorojimo algoritmai, nepriklausomi ištestiniai vietinio apdorojimo algoritmai, apriboti vietinio apdorojimo algoritmai ir globalaus apdorojimo algoritmai. Nepriklausomų taškų algoritmai yra riboto funkcionalumo, kurie atlieka atskirų linijinio elemento viršūnių eliminavimą, neatsižvelgiant į tarpusavio sąryšius, tuo tarpu vietinio apdorojimo algoritmai analizuoja dalies elementų sąryšių charakteristikas, o globalaus apdorojimo algoritmai analizuodami elemento dalis vertina visą elementą iš karto.

Nepriklausomų taškų algoritmai suprastinimo metu neišlaiko jokių matematinių topologinių ryšių tarp kaimyninių elementų viršūnių taškų x , y koordinatų. Šie algoritmai atlikdami linijinio elemento suprastinimą:

- palieka n -ąją linijinio elemento viršūnę, kitas panaikindami;
- atsitiktinai atrenka ir palieka linijinio elemento viršūnes.

2.7 paveiksle pateikta nepriklausomų taškų algoritmų veikimo iliustracija.



2.7 pav. Nepriklausomų taškų algoritmų veikimo iliustracija (pagal Shi, Cheung 2006)

Fig. 2.7. The illustration of Independent Point Algorithms (after Shi, Cheung 2006)

Tokie algoritmai nėra tinkami suprastinant erdvinių duomenų, vaizduojančių gamtinius ir antropogeninius objektus, linijinius elementus, kadangi linijos prastinamos panaikinant viršūnes atsitiktiniu būdu ir todėl elemento padėties tikslumas gali būti gaunamas kiekvieną kartą skirtingas.

Vietinio apdorojimo algoritmuose yra analizuojamos linijinio elemento kaimyninės dvi ar trys viršūnės, kurios pašalinamos arba paliekamos įvertinus:

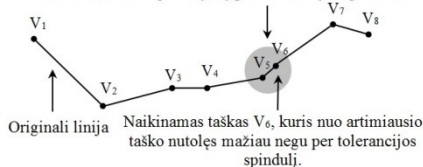
- atstumą tarp kaimyninių viršūnių;
- trikampio, kurį sudaro trys kaimyninės linijinio elemento viršūnės, statmens ilgį, kuris turi būti ne didesnis, nei nustatytos kritinės leistinos tolerancijos reikšmės, t. y. minimalus ilgis arba kampo pokytis.

Šiuose algoritmuose linijinio elemento viršūnės yra panaikinamos, jei atstumas tarp jų ar minėtas statmuo yra mažesnis, nei nustatyta leidžiama tolerancija (2.8 pav.).

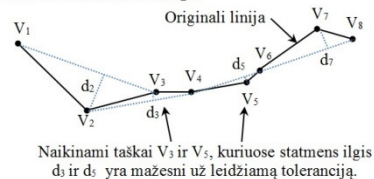
1. Taškų eliminavimas

1.1. Vertinant atstumą tarp kaimyninių viršūnių

Paieškos zona, kurio spindulys lygus tolerancijos spinduliui



1.2. Vertinant statmens ilgį

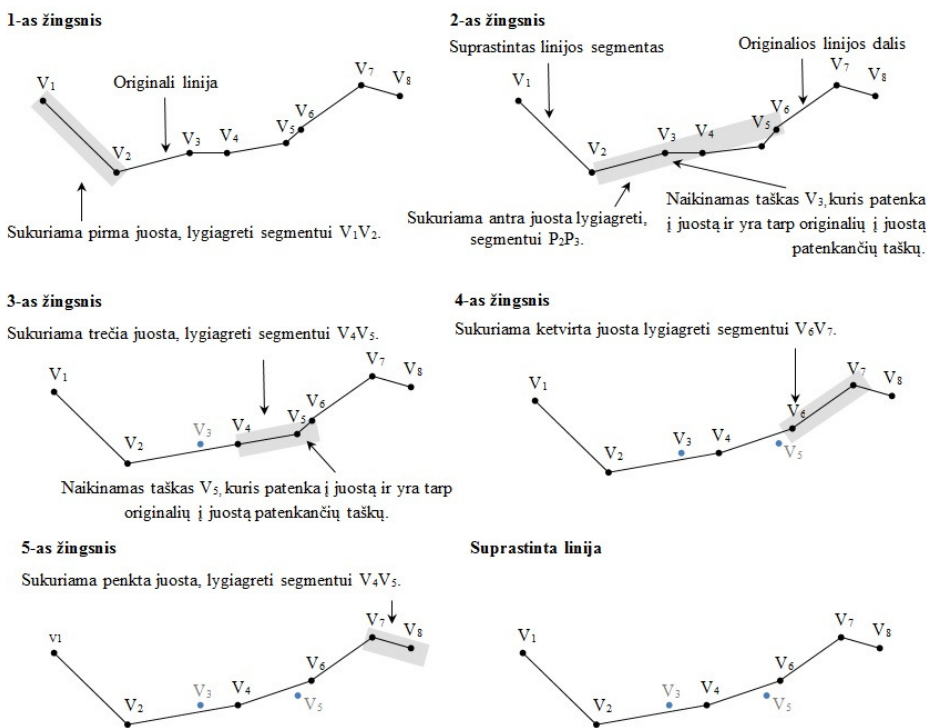


2.8 pav. Vietinio apdorojimo algoritmų veikimo iliustracija (pagal Shi, Cheung 2006)

Fig. 2.8. The illustration of algorithm of local processing (after Shi, Cheung 2006)

Tęsiamieji neapribojami vietinio apdorojimo algoritmai įvertina artimiausias koordinatų poras ir atskirus linijinio elemento segmentus. Šiai kategorijai priklauso Reumann-Witkam algoritmas (Reumann-Witkam 1974) (2.9 pav.). Tokiuose algoritmuose linijiniai elementai yra dalinami į segmentus naudojant nustatyto pločio kuriamas juostas – stačiakampes zonas. Tokia juosta yra brėžiamas lygiagrečiai pirmosios linijinio segmento atkarpos kryptimi, kol tokios juostos kraštinės linija susikerta su linijiniu elementu. Juostoje esantis paskutinis taškas tampa pradžios tašku, kitai brėžiamai paieškos juostai. Juosta

toliau brėžiama lygiagrečiai atkarpai, su kuria susikirto, kryptimi. Šis procesas yra tęsiamas tol, kol į brėžiamą juostą patenka paskutinis prastinamo linijinio elemento taškas – pabaigos taškas. Užbaigus procesą pradinės linijos versijos pradžios ir pabaigos taškai yra grąžinami, senosios viršūnės panaikinamos, o likę viršūnių taškai apjungiami linijomis, taip gaunant suprastintą linijinį elementą.

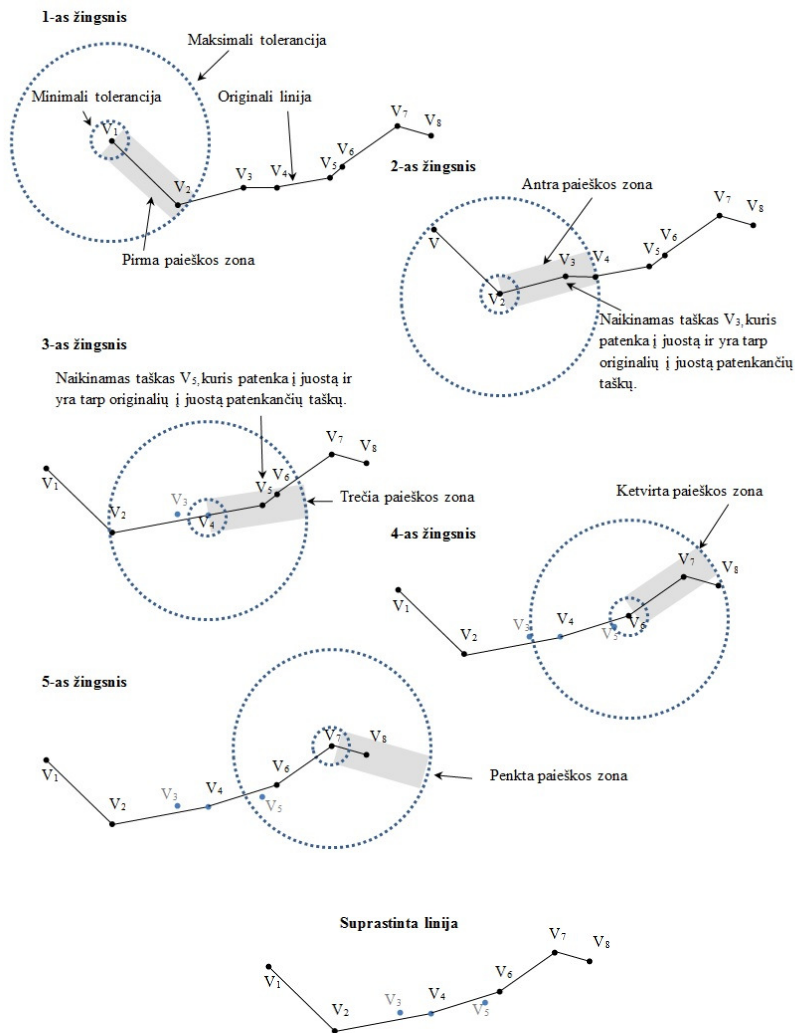


2.9 pav. Reumann-Witkam (1974) supaprastinimo algoritmo veikimo iliustracija (pagal Shi, Cheung 2006)

Fig. 2.9. The illustration of simplification algorithms by Reumann-Witkam (1974) (after Shi, Cheung 2006)

Tęsiamieji apribojami vietinio apdorojimo algoritmai veikia panašiai, kaip ir anksčiau minėtieji tęsiamieji neapribojami vietinio apdorojimo algoritmai, tačiau papildomai yra naudojama linijinio elemento viršūnių analizės (paieškos) zona, kurią apibrėžia nustatyta tolerancija. Opheim supaprastinimo algoritmas (Opheim 1981, 1982) naudoja brėžiamą nustatyto pločio juostą, kaip ir Reumann-Witkam algoritme (Reumann-Witkam 1974), tačiau nustato leidžiamą minimalų ir maksimalų juostos ilgį, kuris priklauso nuo tolerancijos

(2.10 pav.). Šiuo atveju visi taškai patenkantys į režį panaikinami, išskyrus paskutinį režyje esantį tašką.

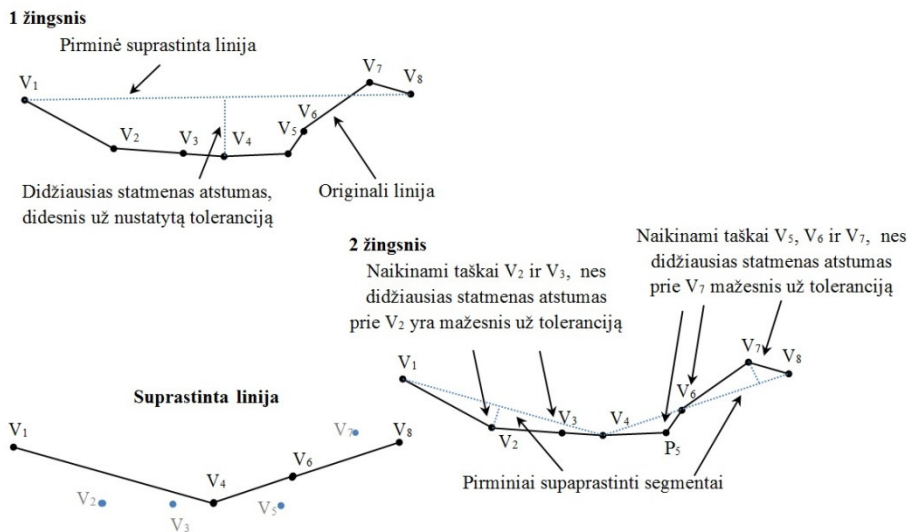


2.10 pav. Opheim suprastinimo algoritmo veikimo iliustracija (1981, 1982)
(pagal Shi, Cheung 2006)

Fig. 2.10. The illustration of the Opheim simplification algorithm (1981, 1982)
(after Shi, Cheung 2006)

Tuo tarpu Lang suprastinimo algoritme (Lang 1969) analizuojamas papildomai trikampio, tarp trijų linijinio elemento viršūnių, statmens ilgis. Jei ilgis yra mažesnis, nei nurodyta tolerancija, viršūnė yra panaikinama taip kaip vietinio apdorojimo algoritmuose vertinant statmens ilgį.

Globalaus apdorojimo algoritmai suprastinimo procese analizuoją visą linijinį elementą. Douglas-Peucker algoritmas (Douglas, Peucker 1973), patikslintas Hersberger ir Snoeynk (1992), naudoja iš anksto nustatytą toleranciją, kuri apibrėžia minimalų leidžiamą statmens atstumą nuo jungiamos linijos atkarpos, siejančias tarpines viršūnes. Linijinio elemento viršūnės, viena kitos atžvilgiu ta pačia kryptimi išsidėsčiusios tolerancijos ribose yra eliminuojamos (2.11 pav.).



2.11 pav. Douglas-Peucker suprastinimo algoritmo veikimo iliustracija (1973) (pagal Shi, Cheung 2006)

Fig. 2.11. The illustration of Douglas-Peucker simplification algorithm (1973) (after Shi, Cheung 2006)

Douglas-Peucker algoritmas yra dažniausiai naudojamas tiek kartografijoje, tiek ir GIS. GIS dažnai naudojamas vingių suprastinimo algoritmas (Wang, Muller 1999) (toliau – Wang algoritmas), kurį panaudojus gaunamas geresnis generalizuotos linijos kartografinis vaizdas. Šis algoritmas analizuoja suprastinamų linijų vingius ir įvertina jų savybes, eliminuodamas nereikšmingiausias iš jų. Algoritme yra taikomos keturios pagrindinės taisyklės:

- nedideli vingiai ir nuokryptai turi būti eliminuojami;
- du vienodi lankai einantys vienas po kito turi būti sujungiami į vieną;
- netiesios linijos neturi būti pakeistos geometriškai tiesiomis linijomis;
- izoliuoti vingiai turi būti padidinti, jei jų dydis yra artimas leidžiamam ribiniam dydžiui.

Atsižvelgiant į teorinių linijinių elementų suprastinimo algoritmų tyrimų rezultatus, tolimesniam eksperimentiniam tyrimui parinkti Douglas-Peuker ir Wang suprastinimo algoritmai, norint įvertinti jų tinkamumą prastinant erdvinį duomenų linijinių elementų ar plotinių elementų ribų konfigūracijas. Taip pat eksperimente naudotas siūlomas praplėstas Wang algoritmas, kuriame po linijinio ar plotinio elemento suprastinimo, taikoma glotninimo operaciją paremtą polinomo apibendrinimu taikant branduolio eksponentę siekiant dar labiau padidinti kartografinį vaizdą, ir papildomai Douglas-Peuker algoritmą, kuris po glotninimo eliminuotų perteklines elemento viršūnes.

Atliekant eksperimentą naudoti GDR10LT erdvinį duomenų elementų klasės „Keliai“ 14337 elementai, vaizduojantys kelių ir gatvių ašines linijas, „Hidro_l“ 8653 elementai, vaizduojantys upių, kanalų ir melioracijos griovių ašines linijas, „Plotai“ 7444 elementai, kurie vaizduoja miškus ir 5237 elementai, vaizduojantys ežerus, tvenkinius ir kūdras. Eksperimente naudotos generalizavimo operacijų tolerancijos reikšmės pateiktos 2.7 lentelėje.

2.7 lentelė. Tolerancijos reikšmės naudotos generalizavimo operacijoms eksperimente
Table 2.7. Value of tolerance used for operations of generalisation in experiment

Generalizavimo algoritmo pavadinimas	Tolerancijos reikšmė, m		
	Wang algoritme	Glottinimo algoritme	Douglas-Peuker algoritme
Wang (praplėstas)	20	20	1
Wang	20	–	–
Douglas-Peuker	–	–	20
Douglas-Peuker	–	–	10
Douglas-Peuker	–	–	5

Siekiant nustatyti tinkamumą eksperimente naudojamiems Wang (praplėstas), Wang ir Douglas-Peuker algoritmams parinktos vienodos tolerancijos reikšmės. Taip pat papildomai atlikti eksperimentai su Douglas-Peuker algoritmu taikant mažesnes tolerancijos reikšmes.

Kadangi erdvinį duomenų generalizavimo procese yra svarbus optimalus elementų kompleksiskumo ir kiekio sumažinimas, todėl vertinant eksperimento rezultatus, gautus taikant skirtingus algoritmus ir tolerancijos reikšmes, atliktas erdvinį duomenų elementų prieš ir po suprastinimo vertinimas, taikant

Podolskaya su kolegomis (2007) patobulintą siūlomą plotinių elementų generalizavimo kokybės įvertinimo metodą, kuris pritaikytas vertinant ir linijinių elementų generalizavimo kokybės vertinime. Šioje analizėje nagrinėjami šie kriterijai:

- viršūnių kiekio pokytis:

$$V_N = \frac{|N_2 - N_1|}{\max(N_1, N_2)} , \quad (2.9)$$

čia N_1 – elemento viršūnių kiekis prieš generalizavimą, N_2 – viršūnių kiekis po generalizavimo. Paprastai atlikus linijinių ir plotinių elementų suprastinimą, sumažėja viršūnių kiekis, todėl $N_1 > N_2$. Tačiau taikant siūlomą Wang praplėstą algoritmą, kuriame naudojama glotninimo operacija, tam, kad padidinti elemento kartografinį vaizdą padidinamas viršūnių kiekis, todėl vertinant praplėstą Wang algoritmą viršūnių kiekio pokyčio vertinti nereikėtų.

- plotų pokytis:

$$V_S = 1 - \frac{|S(p_2) - S(p_1)|}{\max(S(p_1), S(p_2))} , \quad (2.10)$$

čia $S(p_1)$ – elemento plotas prieš generalizavimą, $S(p_2)$ – elemento plotas po generalizavimo;

- perimetro pokytis:

$$V_P = 1 - \frac{|P(p_2) - P(p_1)|}{\max(P(p_1), P(p_2))} , \quad (2.11)$$

čia $P(p_1)$ – elemento perimetras prieš generalizavimą, $P(p_2)$ – elemento perimetras po generalizavimo;

- elementų konfigūracijos pokytis, kuriai išreikšti naudojama posūkio kampo funkcija (Frank, Ester 2006):

$$V_{TF} = 1 - \frac{S(TF_1 \Delta TF_2)}{\max(S(TF_1), S(TF_2))} , \quad (2.12)$$

čia $S(TF_1)$ – plotas, kurį sudaro elemento prieš generalizavimą konfigūracijos posūkio funkcija su x ašimi, $S(TF_2)$ – plotas, kurį sudaro elemento po generalizavimo konfigūracijos posūkio funkcija su x ašimi, $S(TF_1 \Delta TF_2)$ – ploto skirtumas tarp posūkio $S(TF_1)$ ir $S(TF_2)$.

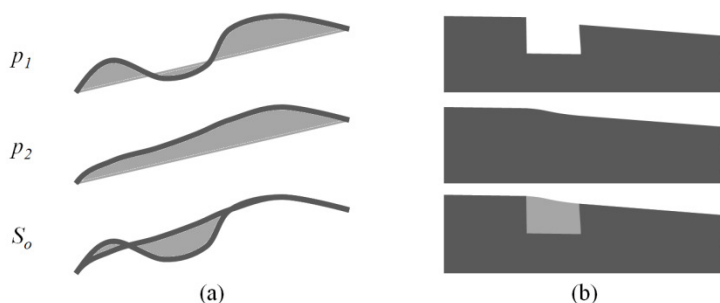
Nagrinėjamu atveju yra analizuojami tik suprastintos konfigūracijos plotiniai ir linijiniai elementai, kuriems nebuvo taikomos jokios kitos

generalizavimo operacijos, tokios kaip perstūmimas, apjungimas ir panašiai. Todėl elementai po suprastinimo radikaliai nepasikeitė. Dėl šios priežasties, kaip pavaizduota 2.12 paveiksle, linijinių ir plotinių elementų konfigūraciją ir jų skirtumą prieš ir po generalizavimo galima įvertinti plotu. Analizėje naudota paprastesnė formulė, kuri tinkama konfigūracijos pokyčiui vertinti:

$$V_0 = 1 - \frac{S_0}{\max(P(p_1), P(p_2))}, \quad (2.13)$$

čia S_0 – elemento prieš ir po generalizavimo dengiamo ploto skirtumas, $P(p_1)$ – elemento plotas prieš generalizavimą, $P(p_2)$ – elemento plotas po generalizavimo.

Paveiksle 2.12 pavaizduota (2.13) formulėje naudojamų sutapdinimo vertinimo parametrų atvaizdavimas.



2.12 pav. Formulėje (2.13) naudojamų sutapimo vertinimo parametrų atvaizdavimas
Fig. 2.12. The visualisation of parameters of matching evaluation use in (2.13) formula

Bendra generalizuotų linijinių ir plotinių elementų kokybės išraiška skaičiuojama naudojant svorines V_N , V_S , V_P ir V_O reikšmes:

$$V = c_0 \cdot V_0 + c_S \cdot V_S + c_P \cdot V_P + c_N \cdot V_N, \quad (2.14)$$

čia $c_0 + c_S + c_P + c_N = 1$, kur c_0 , c_S , c_P , c_N yra skirtingos V_0 , V_N , V_S , V_P svorių reikšmės.

Atliekant eksperimentą formulėje (2.14) taikytos svorių reikšmės: $c_N = 0,5$, $c_S = 0,167$, $c_P = 0,167$, $c_0 = 0,167$. Tokios reikšmės leidžia tolygiai įvertinti elemento konfigūracijos suprastinimą (V_N) ir panašumą (V_0 , V_S , V_P) (Podolskaya *et al.* 2007). Tačiau, vertinant praplėstą Wang algoritmą nebuvo vertinama V_N pagal (2.9) formulę, kadangi glotninimo algoritmai padidina elemento viršūnių kiekį, tam kad gauti geresnį kartografinį vaizdą. Todėl taikant šį algoritmą

viršūnių kiekio įvertinimas neatspindės konfigūracijos suprastinimo, nors lyginant su pirmine elemento konfigūracija bus pakitusi ir paprastesnė.

Apibendrinti eksperimento rezultatai pateikti 2.8 lentelėje.

2.8 lentelė. Tolerancijos reikšmės naudotos generalizavimo operacijoms eksperimente
Table 2.8. Value of tolerance used for operations of generalisation in experiment

Erdvinių duomenų elementų klasė	Praplėstas Wang (tolerancija 20 m)		Wang (tolerancija 20 m)		Douglas-Peucker (tolerancija 10 m)		Douglas-Peucker (tolerancija 20 m)		Douglas-Peucker (tolerancija 5 m)	
HIDRO_L	Maks:	1	Maks:	1	Maks:	1	Maks:	1	Maks:	1
	Min.	0,125	Min.	0,088	Min.	0,17	Min.	0,17	Min.	0,17
	Vid.	0,816	Vid.	0,925	Vid.	0,727	Vid.	0,712	Vid.	0,746
PLOTAI (ežerai, tvenkiniai, kūdros)	Maks:	0,998	Maks:	1	Maks:	0,745	Maks:	0,714	Maks:	1
	Min.	0,031	Min.	0,039	Min.	0,024	Min.	0,024	Min.	0,024
	Vid.	0,7	Vid.	0,925	Vid.	0,44	Vid.	0,407	Vid.	0,519
KELIAI	Maks:	1	Maks:	1	Maks:	1	Maks:	1	Maks:	1
	Min.	0,19	Min.	0,109	Min.	0,304	Min.	0,296	Min.	0,333
	Vid.	0,748	Vid.	0,921	Vid.	0,642	Vid.	0,628	Vid.	0,662
PLOTAI (miškai)	Maks:	0,998	Maks:	1	Maks:	1	Maks:	1	Maks:	1
	Min.	0,003	Min.	0,003	Min.	0,002	Min.	0,002	Min.	0,002
	Vid.	0,758	Vid.	0,906	Vid.	0,561	Vid.	0,512	Vid.	0,61

Atlikus eksperimentinius tyrimus, nustatyta, kad erdviniams duomenims taikant Wang algoritmą, bendra suprastintų elementų vidutinė kokybės išraiška yra apie 0,9, praplėstą Wang algoritmą – 0,7, o taikant Douglas-Peucker algoritmą 0,56, kai tolerancijos reikšmė yra 20 m, 0,59, kai 10 m ir 0,63, kai 5 m. Atsižvelgus į tyrimų rezultatus, siūloma pirmuosius su algoritmus taikyti gamtinių objektų generalizavimo modeliuose, o Douglas-Peucker antropogeninių objektų suprastinime, atitinkamai parinkus tinkamą tolerancijos reikšmę, optimaliai išlaikant suprastinamų elementų kokybę.

Tačiau atliekant erdvinių duomenų generalizavimą, šiuos algoritmus reikia kombinuoti su kitomis erdvinių duomenų modeliavimo funkcijomis, kadangi suprastinus plotinius elementus jie gali dengti vienas kitą arba tarp kaimyninių elementų gali atsirasti bereikalingi tarpai. Tai atsitinka dėl to, kad šie algoritmai neatsižvelgia į ryšius tarp kaimyninių elementų. To galima išvengti plotinius elementus pirmiausiai transformavus į linijas, kurias suprastinus, vėl sukuriama plotiniai elementai ir iš naujo priskiriamos anksčiau buvusios erdvinių duomenų elementų savybių reikšmės. Tokį erdvinių elementų suprastinimo metodą siūlo Nielsen ir Meyer (2007), kuris yra taikomas Olandijoje generalizuojant

TOP10NL erdvinius duomenis, skirtus TOP50NL sukūrimui. Algoritmas aprašytas 2.9 lentelėje.

2.9 lentelė. Erdvinių duomenų plotinių elementų supaprastinimo algoritmas

Table 2.9. Algorithm of simplification of spatial data polygons

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Įkeliama erdvinių duomenų plotinių elementų klasė (<i>EK1_plotiniai</i>).	—
2.	<i>EK1_plotiniai</i> elementai transformuojami į taškinius elementus, išlaikant plotiniams elementams priskirtas savybių reikšmes. Rezultatas: erdvinių duomenų elementų klasė <i>EK1_taşkiniai</i> .	<i>Feature to Point</i>
3.	<i>EK1_plotiniai</i> elementai transformuojami į linijinius elementus, neišlaikant plotiniams elementams priskirtų savybių reikšmių. Rezultatas: erdvinių duomenų elementų klasė <i>EK1_linijiniai</i> .	<i>Feature to Line</i>
4.	<i>EK1_linijiniai</i> elementai supaprastinami. Rezultatas: erdvinių duomenų elementų klasė <i>EK2_linijiniai</i> , kurioje saugomi supaprastinti elementai.	<i>Simplify Lines</i>
5.	Iš elementų klasėje <i>EK2_linijiniai</i> elementų sukuriama plotinių elementų klasė. Rezultatas: elementų klasė <i>EK2_plotiniai</i> .	<i>Feature to Polygon</i>
6.	Penktame žingsnyje sukurtiems elementams priskiriamos savybių reikšmės taikytos ankstesniems plotiniams elementams saugomiems elementų klasėje <i>EK1_plotiniai</i> . Šiuo atveju <i>EK2_plotiniai</i> elementams priskiriamos savybių reikšmės <i>EK1_taşkiniai</i> elementų, kurie kerta pastaruosius.	<i>Spatial Join</i> <i>Calculate Field</i>
	Rezultatas: elementų klasė <i>EK2_plotiniai</i> , kurioje saugomi supaprastinti elementai.	—

Šio algoritmo privalumas yra toks, kad atlikus elementų supaprastinimą, neatsiranda persidengimų ar nereikalingų tarpų tarp elementų. Bet pagrindinis trūkumas yra tai, kad visiems elementams gali būti taikomas tik vienas ir tas pats supaprastinimo algoritmas, taikant tuos pačius generalizavimo parametrus (pavyzdžiui, leidžiamą toleranciją). Iš dalies to išvengti galima, linijinius elementus išskaidžius į atskiras elementų klases, priklausomai nuo to, kokius supaprastinimo algoritmus ar supaprastinimo parametrus planuojama naudoti. Tai įgyvendinant, 2.9 lentelėje aprašyto algoritmo trečiame žingsnyje reikia sukurtiems linijiniams elementams priskirti plotinių elementų savybių reikšmes pagal prioritetą. Atlikus linijinių elementų supaprastinimą pagal skirtingus parametrus, elementai apjungiami į vieną elementų klasę. Toliau vykdomas 2.9 lentelėje aprašyto algoritmo penktas žingsnis.

Taip pat norint taikyti skirtingus parametrus ar algoritmus plotiniams elementams priklausomai nuo jų savybių, galima taikyti siūlomą algoritmą aprašytą 2.10 lentelėje, kuriame galima suprastinti tiek plotinius, tiek linijinius elementus.

2.10 lentelė. Erdvinių duomenų plotinių elementų suprastinimo algoritmas

Table 2.10. Algorithm of simplification of spatial data polygons

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Įkeliama erdvinių duomenų plotinių elementų klasė <i>EK1_plotiniai</i> .	–
2.	Atliekama elementų atranka pagal elementams priskirtų savybių reikšmes, atitinkamai perkeltiant atrinktus elementus į atskiras elementų klases <i>EK2_plotiniai</i> .	<i>Feature Class to Feature Class</i>
3.	Atskirose elementų klasėse <i>EK2_plotiniai</i> elementai suprastinami naudojant skirtingus elementų suprastinimo algoritmus ir (ar) parametrus.	<i>Simplify Polygons</i> arba <i>Simplify Building</i> ,
	Šiame žingsnyje atskiroms elementų klasėms gali būti naudojamas suprastinimo algoritmas aprašytas 2.7 lentelėje. Rezultatas: elementų klasės <i>EK3_plotiniai</i> , kuriose saugomi suprastinti plotiniai elementai.	<i>Repair Geometry</i> .
4.	Elementų klasių <i>EK3_plotiniai</i> elementai sujungiami į vieną elementų klasę <i>EK4_plotiniai</i> taip, kad persidengimo vietose būtų sukurti nauji plotiniai elementai, kuriems būtų priskirtos abiejų persidengiančių elementų savybių reikšmės.	<i>Union</i>
5.	Elementų klasėje <i>EK4_plotiniai</i> elementams, turintiems dviejų elementų, kurie persidengė, savybių reikšmes, yra paliekama to elemento savybės reikšmė, kuri pagal prioritetą yra svarbesnė.	<i>Select by Attribute</i> <i>Calculate Field</i>
6.	Elementų klasėje <i>EK4_plotiniai</i> elementai apjungiami atsižvelgiant į jiems priskirtas unikalias savybių reikšmes.	<i>Dissolve</i>
7.	Taikant šį algoritmą gali likti tarpai tarp plotinių elementų, todėl šiame žingsnyje tokiuose tarpuose turi būti sukuriama plotiniai elementai.	<i>Create Topology</i> , <i>Add Feature Class</i> , <i>Add Rule (Must Not Have Gaps)</i> , <i>Create Features</i> .
8.	Septintame žingsnyje gauti plotiniai elementai apjungiami prie tų kaimyninių elementų, su kuriais bendra kraštinė yra ilgiausia.	<i>Eliminate</i>
	Rezultatas: elementų klasė <i>EK4_plotiniai</i> , kurioje saugomi supaprastinti elementai.	–

Taikant šį algoritmą, erdvinių duomenų elementus išskaidžius į atskiras klases, galima taikyti skirtingus elementų suprastinimo algoritmus ir juose taikomus parametrus. Vėliau apjungus šiuos duomenis, persidengimai yra eliminuojami, juos prijungiant prioriteto tvarka prie kaimyninių elementų.

2.2.4. Linijinių elementų atšakų eliminavimas

Algoritmas yra skirtas linijinių elementų, sudarančių geometrinį tinklą, atšakų (angl. *dangle*) identifikavimui ir panaikinimui atsižvelgiant į jo prioretiškumą reikšmę.

Savino (2011) pasiūlė sprendimą kelių tinklo atšakų eliminavimui, kuris turi būti vykdomas keliomis iteracijomis, kadangi panaikinus atšaką, kurios ilgis yra mažesnis, nei leistinas linijinio elemento ilgis, gali atsirasti nauja atšaka, kurios ilgis vėl bus mažesnis, nei leistinas minimumas. Todėl tokios tinklo atšakos yra naikinamos tol, kol eliminuotų linijinių elementų, kurių sutampa pradžios ir (ar) pabaigos taškų x , y koordinatės, bendras ilgis yra mažesnis, nei leistinas minimalus ilgis.

Siūlomas patobulintas algoritmas pateiktas 2.11 lentelėje.

2.11 lentelė. Atšakų eliminavimo algoritmas

Table 2.11. Algorithm of elimination of dangles

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Įkeliama erdvinių duomenų plotinių elementų klasė <i>EK1_plotiniai</i> .	—
2.	<i>EK1_plotiniai</i> elementams priskiriamos prioriteto reikšmės, vadovaujantis elementams priskirtomis savybių reikšmėmis ir (ar) sąsaja su kitais elementais. Elementų atitinkančių tam tikrą prioritetą atrankai galima naudoti 2.2.1 poskyryje pateikiamus algoritmus.	<i>Select by Attribute</i> , <i>Calculate Field</i> .
3.	Analizuojama elementų klasė <i>EK1_linijiniai</i> . Sukuriami taškiniai elementai <i>EK1_linijiniai</i> elementų pradžios arba pabaigos taškuose, jei jie neturi bendros pradžios ar pabaigos su kitu linijiniu elementu <i>EK1_linijiniai</i> elementų klasėje. Rezultatas: erdvinių duomenų elementų klasė <i>EK1_taškiniai</i> .	<i>Feature Vertices to Point (dangle)</i>
4.	Atrenkami elementai, esantys <i>EK1_linijiniai</i> , kurie liečia elementus <i>EK1_taškiniai</i> , t. y. atrenkamos linijos, kurių pradžia ir (arba) pabaiga turi bendrą pradžią ar pabaigą su <i>EK1_taškiniai</i> elementu.	<i>Select Features by Location</i> , <i>Calculate Field</i> .

2.9 lentelės pabaiga

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
	Iš atrankos pašalinami elementai, kurių ilgis yra didesnis, nei nustatytas leidžiamas minimalus atšakos ilgis. Taip pat pašalinami elementai, turintys didesnę prioriteto reikšmę, jei atšakų ilgiai yra vienodi. Atrinktiems elementams priskiriama reikšmė „atšaka“.	
5.	Analizuojami <i>EK1_linijiniai</i> elementai, turintys savybės reikšmę „atšaka“. Elementai turintys žemesnę prioriteto reikšmę, panaikinami.	<i>Select by Attributes, Deletes.</i>
6.	Algoritmo antras, trečias ir ketvirtas žingsniai iš eilės kartojamas keliomis iteracijomis, tol kol eliminuojamos visos mažesnę prioritetą turinčios atšakos, kurių ilgis yra mažesnis, nei nustatytas minimumas.	–
	Rezultatas: elementų klasė <i>EK1_linijiniai</i> , kurioje saugomi linijinių elementų tinklas, neturintis perteklinių mažo prioriteto atšakų, kurių ilgis yra mažesnis, nei nustatytas leidžiamas minimumas.	–

Lyginant su Savino (2011) siūlomu algoritmu 2.11 lentelėje pateiktas algoritmas suranda ne tik elementus, kurių ilgis yra mažesnis, nei leidžiamas minimumas, bet ir įvertina jų svarbumą visame linijų tinkle, ir prireikus juos panaikina. Šiame algoritme reikalinga procesus vykdyti taip pat keliomis iteracijomis. Algoritmas kartojamas pradedant ieškoti trumpiausių atšakų tol, kol surandamas pasirinkto dydžio atšakos ilgis. Šiuo atveju išeities ir galutiniai elementai turi būti linijiniai. Būtina sąlyga, kad erdvinių duomenų elementams turi būti suteiktas prioritetas, tam kad nebūtų išnaikinti svarbūs elementai. Prioritetai turi būti priskiriami priklausomai nuo erdvinių duomenų specifiškumo ir galutinio erdvinių duomenų rinkinio modelio.

2.2.5. Plotinio elemento centrinės linijos kūrimas

Sprendimai, skirti sukurti linijinį elementą identifikuojantį plotinio elemento centrinę liniją yra nagrinėjami keleto mokslininkų (Ai, Oosterom 2002; Aichholzer *et al.* 1997; Chithambaram *et al.* 1991; Bader, Weibel 1997; Roberts *et al.* 2005; Tanase, Veltkamp 2004).

Autorė siūlo centrinės linijos sukūrimo algoritme naudoti Voronojaus diagramą. Tokį principą panaudojo Roberts su kolegomis (2005) algoritme, kuriame sukuriama plotinių elementų, vaizduojančių gatves, ašinės linijos. Autorė tą patį principą pritaikė kuriant centrinės linijas bet kokios konfigūracijos plotiniams elementams.

Siūlomas algoritmas sukuria Voronojaus diagramą tarp plotinių elementų kraštinių viršūnių taškų. Nustatomi ir sukuriami centro taškai, per kuriuos išvedama centrinė linija. Siūlomo algoritmo apibendrinta seka pateikta 2.12 lentelėje.

2.12 lentelė. Plotinių elementų centrinės linijos sukūrimo algoritmas

Table 2.12. Algorithm of creating centre line of polygon feature

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Įkeliama erdvinių duomenų plotinių elementų klasė <i>EK1_plotiniai</i> .	—
2.	Įkeliama erdvinių duomenų linijinių elementų klasė <i>EK1_plotiniai</i> , kurios elementai bus naudojami sukuriant jų centrinės linijas vaizduojančius linijinius elementus.	—
3.	Sutankinamos <i>EK1_plotiniai</i> elementų kraštinių viršūnės nustatytu atstumu. Kuo atstumas bus mažesnis, tuo tikslesnė bus sukurta plotinio elemento centrinė linija.	<i>Densify</i>
4.	Naudojant <i>EK1_plotiniai</i> elementų kraštinės viršūnes sukuriami taškai. Rezultatas: elementų klasė <i>EK1_taškiniai</i> .	<i>Feature Vertices to Point</i>
5.	Naudojant elementų klasės <i>EK1_taškiniai</i> elementus, sukurama Voronojaus diagrama. Rezultatas: elementų klasė <i>EK2_plotiniai</i> .	<i>Create Thiessen Polygons</i>
6.	Naudojant elementų klasės <i>EK1_plotiniai</i> elementus, „apkerpami“ <i>EK2_plotiniai</i> elementai, kurie vėliau apjungiami su <i>EK1_plotiniai</i> elementais. Rezultatas: elementų klasė <i>EK3_plotiniai</i> .	<i>Clip, Union.</i>
7.	Naudojant <i>EK3_plotiniai</i> elementus sukuriami linijiniai elementai. Rezultatas: elementų klasė <i>EK2_linijiniai</i> .	<i>Feature to Lines</i>
8.	Atrenkami <i>EK2_linijiniai</i> elementai, kurie liečia plotinių elementų klasės <i>EK1_plotiniai</i> elementų kraštines. Atrinkti elementai panaikinami. Rezultatas: elementų klasė <i>EK2_linijiniai</i> , kurioje saugomi linijiniai elementai, esantys plotinio elemento centre.	<i>Select by Location, Delete.</i>
9.	Priklausomai nuo plotinio elemento konfigūracijos, <i>EK2_linijiniai</i> elementai gali turėti trūkius, todėl rekomenduojama <i>EK2_linijiniai</i> elementus pratęsti nustatytu atstumu tiese iki kaimyninio elemento. Panaikinamos <i>EK2_linijiniai</i> elementų atšakos. Taikomas algoritmas aprašytas 2.9 lentelėje.	<i>Extend Line</i> arba <i>Trim</i>

2.12 lentelės pabaiga

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
10.	Apjungiami <i>EK2_linijiniai</i> elementai, kurių pradžios ir (ar) pabaigos taškų x, y koordinatės sutampa su kito elemento pradžios ir (ar) pabaigos taškų x, y koordinatėmis.	<i>Dissolve</i>
11.	<i>EK2_linijiniai</i> elementai suprastinami, taikant <i>Peucker-Douglas</i> suprastinimo algoritmą. Rezultatas: elementų klasė <i>EK3_linijiniai</i> .	<i>Simplify Line</i>
12.	<i>EK3_linijiniai</i> elementai pritraukiami prie <i>EK1_plotiniai</i> elementai kraštinių.	<i>Extend Line</i> arba <i>Trim Line</i>
	Rezultatas: elementų klasė <i>EK3_linijiniai</i> , kurioje saugomos sukurtos <i>EK1_plotiniai</i> elementų centrinės linijos.	

2.12 lentelėje pateiktas algoritmas sukuria plotinių elementų centrinės linijas, tačiau priklausomai nuo plotinio elemento konfigūracijos ne visuomet yra elemento centre. Taip pat atkeiptinas dėmesys, kad norint gauti vieną liniją vaizduojančią plotinio elemento centrą reikalinga eliminuoti visas viduje plotinio elemento esančias kiaurymes, kadangi šiuo atveju gausis kelios išsišakojančios linijos, kurios gali susijungti.

2.2.6. Linijinių ir plotinių elementų sutapdinimas

Linijinių ar plotinių elementų kraštinių sutapdinimas su kitos erdvių duomenų elementų klasės elementais yra aktualus uždavinys generalizuojant kelias elementų klases. Pavyzdžiui, plotinių elementų, vaizduojančių žemės dangą, riba turi sutapti su linijiniais elementais vaizduojančiais kelių ar upių ašines linijas. Jones su kolegomis (1995), Savino (2011) ir kiti mokslininkai siūlo naudoti Delaunė trianguliaciją. Tačiau atliekant tyrimus naudojant GDR10LT erdvinius duomenis nustatytos dvi pagrindinės problemos: perkėlus kraštinę vis tiek gali likti trumpos sutapdinamo elemento kraštinės atkarpos, kurios nebus pritrauktos; tokiu procesui atlikti reikalingos didelės laiko sąnaudos (Papšienė *et al.* 2013a). Todėl siūloma šį uždavinį spręsti ne fiziškai perkeliant elemento viršūnes prie elemento kraštinės su kuriuo jis turi būti sutapdinamas, bet jo kraštinę pakeičiant to elemento, su kuriuo turi būti sutapdinama, kraštine. Realizuojant šį sprendimą galimi du būdai. Jei plotiniai elementai, dengiantys visą teritoriją, turi būti sutapdinami su linijiniais elementais, plotinius elementus reikia eksportuoti į linijinius elementus, į juos integruoti linijinius elementus, su kuriais turi būti sutapdinama. Iš gautų linijinių elementų sukuriama plotiniai elementai, jiems priskiriamos ankstesnių savybių reikšmės. Maži plotiniai

elementai yra eliminuojami prie kaimyninių elementų, su kuriais jie turi ilgiausią bendrą kraštinę. Šiuo atveju būtina eliminuojant atsižvelgti, kad eliminuojamas elementas, nebūtų prijungtas prie kaimyninio elemento kitoje linijinio elemento pusėje. Tačiau šis būdas yra efektyvus, tik kai linijinis elementas, prie kurio reikia priderinti plotinį elementą jį kerta. Tokiais atvejais, kai linijinis elementas nekerta plotinio elemento arba prie jo turi būti priderintas kitas linijinis elementas, arba, kai linijinis elementas turi būti priderintas prie plotinio elemento, siūloma naudoti algoritmą aprašytą 2.13 lentelėje.

2.13 lentelė. Elementų kraštinių sutapdinimo algoritmas

Table 2.13. Algorithm of matching boundaries of features

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Įkeliami erdvinių duomenų elementų klasės <i>EK1</i> ir <i>EK_derinti</i> , kurių <i>EK1 elementų</i> kraštinės turi būti sutapdintos su elementų klasės <i>EK_derinti</i> elementais.	–
2.	Aplink <i>EK1</i> elementus sukuriamas buferis, kurio atstumo reikšmė yra lygi nustatytai atstumo reikšmei, kuriai esant tarp analizuojamų elementų jie turi būti sutapdinti. Rezultatas: elementų klasė <i>EK1_buferis</i> .	<i>Buffer</i>
3.	Iškerpami <i>EK_derinti</i> elementai, kurie patenka į <i>EK1_bufferis</i> . Iškirpti elementai apjungiami. Rezultatas: elementų klasė <i>EK_derinti1</i> .	<i>Clip, Dissolve</i>
4.	Papildomai reikia įvertinti gautų elementų ilgį. Atrenkami <i>EK_derinti1</i> elementai, kurių ilgis yra trumpesnis, nei nustatytas minimumas. Atrinkti elementai ištrinami. Rezultatas: elementų klasė <i>EK_derinti1</i> , kurioje elementų ilgis yra didesnis, nei nustatytas minimumas. Šios linijos bus integruojamos į galutinę suderintų elementų klasę.	<i>Select by Attribute, Delete</i>
5.	Aplink <i>EK_derinti1</i> elementus sukuriamas buferis, kurio atstumo reikšmė yra lygi nustatytai atstumo reikšmei, kuriai esant tarp analizuojamų elementų jie turi būti sutapdinti. Rezultatas: elementų klasė <i>EK_derinti1_buferis</i> .	
6.	Iškerpami <i>EK1</i> elementai, kurie patenka į <i>EK_derinti1_buferis</i> . Iškirpti elementai pašalinami.	<i>Clip,</i>
7.	Į elementų klasę <i>EK1</i> įkeliami elementų klasės <i>EK_derinti1</i> elementai.	<i>Append</i>

2.13 lentelės pabaiga

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
	Kadangi tarp likusių elementų pradžios ir (ar) pabaigos taškų x, y koordinatės dažniausiai nesutaps su <i>EK_derinti1</i> elementų pradžios ir (ar) pabaigos taškų koordinatėmis, reikia sukurti linijas tarp tokių taškų.	
8.	Atliekama dviejų artimiausių taškinių elementų paieška tarp elementų <i>EK1</i> elementų. Tai reikalinga dėl to, kad ieškoma toje pačioje elementų klasėje. Šiuo atveju pirmasis artimiausias elementas visuomet bus jis pats, o antrasis – nutolęs nuo jo tam tikru atstumu. Rezultatas: lentelė, kurioje nurodytos artimiausių elementų poros.	<i>Generate Near Table</i>
9.	<i>EK1</i> elementai apjungiami su lentele, gauta penktame žingsnyje, naudojant identifikatorius, kurie unikaliai identifikuoja <i>EK1</i> elementus ir elementus minėtoje lentelėje. Sukuriamos linijos tarp rastų kaimyninių elementų, kurios išsaugomos elementų klasėje <i>EK2</i> .	<i>Add Join, Add XY, Calculate, Create line from XY.</i>
10.	<i>EK2</i> elementai integruojami į <i>EK1</i> elementų klasę. <i>EK1</i> elementai apjungiami.	<i>Append, Dissolve</i>
	Rezultatas: elementų klasė <i>EK1</i> , kurioje saugomos linijos sutapdintos su <i>EK_derinti</i> elementais.	–

2.13 lentelėje aprašytame algoritme naudojami tik linijiniai elementai, todėl jei sutapdinime naudojami plotiniai elementai, juos reikia transformuoti į linijinius elementus. Vėliau, sutapdinus elementus, plotiniai elementai sukuriama iš naujo iš linijinių elementų, jiems priskiriant ankstesnių plotinių elementų savybių reikšmes.

2.3. Erdvinių duomenų pokyčių nustatymo ir įvertinimo metodai

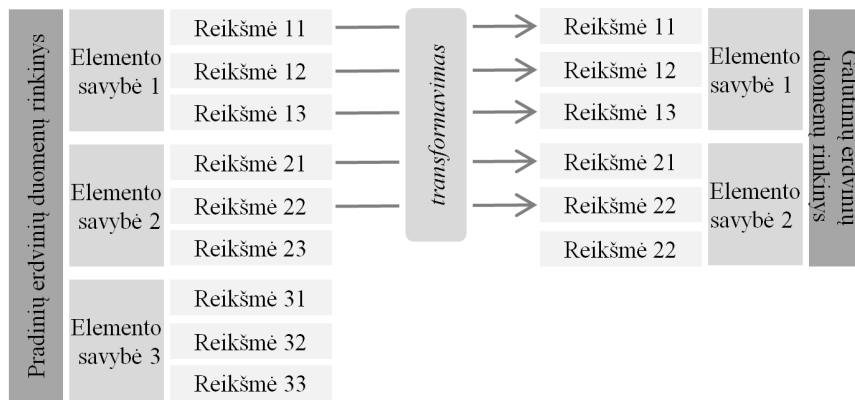
Erdvinių duomenų harmonizavimo apimtis priklauso nuo poreikio. Pavyzdžiui, INSPIRE apimtyje atliekant erdvinių duomenų harmonizavimą turi būti kiekvieną kartą vykdomas visų erdvinių duomenų transformavimas. Tačiau ne visais atvejais yra svarbu transformuoti visus duomenis, ypač tuomet kai pokyčiai pradiniam erdvinių duomenų rinkinyje iš principo nepakeis galutinio erdvinių duomenų rinkinio atlikus harmonizavimą. Todėl vienas iš disertacijos darbo uždavinių yra patobulinti erdvinių duomenų pokyčių nustatymo ir

įvertinimo metodus, kurie leistų pakartotiniame harmonizavimo procese ne tik perduoti informaciją apie pasikeitusį elementą, bet ir nustatyti pradinių erdvinių duomenų rinkinio elementų pokyčius ir jų įtaką harmonizavimo procesui. Tai yra nustačius pokyčius pradiname erdvinių duomenų rinkinyje, jie turi būti įvertinti ir išmatuoti, norint nustatyti, ar bus matomas pokytis galutiniame erdvinių duomenų rinkinyje atlikus harmonizavimą. Todėl šiame poskyryje siūlomi metodai, skirti erdvinių duomenų pokyčių nustatymui ir nustatytų pokyčių įvertinimui.

2.3.1. Erdvinių duomenų pokyčių nustatymas cikliniu perteklinės informacijos tikrinimo metodu

Erdvinių duomenų rinkinio elementų pokyčiams nustatyti siūlomas patobulintas Li ir Fei (2008) aprašytas pokyčių nustatymo metodas. Siūlomas metodas remiasi ciklinės perteklinės informacijos tikrinimo (angl. *cyclic redundancy check, CRC*) principais. CRC yra klaidų nustatymo kodas (Peterson, Brown 1961), kuris plačiai naudojamas kompiuteriniuose tinkluose ir magnetiniuose įrenginiuose, norint patikrinti, ar įrašyti duomenys yra pilnos apimties. Šis siūlomas algoritmas apima ciklišką patikros procesą, kurio metu vertinamas informacijos perteklius ar trūkumas. Siūlomame metode CRC procedūra skirta nepasikeitusių, pasikeitusių, naujų bei ištrintų erdvinių duomenų elementų nustatymui. Vertinant pokyčius, nustatomas kiekvieno erdvinių duomenų elemento identifikatorius (toliau – CRC identifikatorius), naudojantis elementams priskirtas savybių ir geometrijos parametrų reikšmės.

Norint nustatyti erdvinių duomenų elementui priskirtų savybių pokytį, CRC identifikatorių reikia sudaryti iš erdvinių duomenų elementams priskirtų savybių reikšmių. Taip pat tokį principą siūlo Li ir Fei (2008). Tačiau parenkant savybes, kurių reikšmės sudarytų identifikatorių, reikia atsižvelgti į tai, kurios iš jų yra naudojamos vykdant transformacijas harmonizavimo procese. Identifikatorių turi sudaryti erdvinių duomenų elementų savybės, kurių visos arba dalis reikšmių yra naudojamos transformavimo procese vykdant erdvinių duomenų harmonizavimą. Priešingu atveju pokytis neturės jokios įtakos tolimesniam harmonizavimo procesui. Įtakos gali neturėti ir tų savybių pokytis, kurių tik dalis reikšmių naudojama vykdant transformaciją. Tačiau šiuo atveju būtina jas naudoti, kadangi priskirta nauja reikšmė gali būti naudojama vėliau vykdant transformavimą. Savybių, kurios turi sudaryti identifikatorių principinė parinkimo schema pavaizduota 2.13 paveiksle.



$CRC_ID = \text{„Elemento savybė 1“} \& \text{„Elemento savybė 2“}$

2.13 pav. Principinė erdviųjų duomenų elemento savybių parinkimo schema identifikatoriui sudaryti

Fig. 2.13. Common scheme of selection of attributes of spatial data elements for composing the identifier

Kaip pavaizduota 2.13 paveiksle, sudarant identifikatorių reikės naudoti „Elemento savybė 1“ ir „Elemento savybė 2“ reikšmes, kadangi šių erdviųjų duomenų elementams priskirtų savybių reikšmės bus naudojamos transformavimo procese. Tačiau vien elemento savybių reikšmių sudarant CRC identifikatorių neužtenka, kadangi paprastai savybių reikšmės nėra unikalios. Todėl šiuo atveju toks CRC identifikatorius neleis unikalčiai identifikuoti objekto. Todėl papildomai turi būti naudojamos:

- erdviųjų duomenų elementų reikšmės, unikalčiai identifikuojančios kiekvieną elementą. Pavyzdžiui, GDR10LT kiekvienas elementas turi unikalų identifikatorių, kuris išlieka unikalus visame erdviųjų duomenų rinkinyje visą elemento gyvavimo laiką;
- CRC identifikatorių, skirtą elementų konfigūracijos ar vietos pokyčiui nustatyti.

Norint nustatyti erdviųjų duomenų elementų konfigūracijos ar vietos pokytį, CRC identifikatorių reikia sudaryti iš elemento geometriją apibūdinančių parametrų reikšmių.

Erdviųjų duomenų elemento geometrijos pokytį parodantys parametrai pateikti 2.14 lentelėje. Tačiau ne visais atvejais šioje lentelėje nurodytų pavienių parametrų reikšmių vertinimas gali reikšti elemento konfigūracijos arba jo vietos pokytį, išskyrus tuos atvejus, kai identifikatorius sudaromas iš kiekvienos taškinio, linijinio ar plotinio elemento viršūnės reikšmės. Li ir Fei (2008) pokyčiams nustatyti siūlo naudoti erdviųjų duomenų unikalčius identifikatorius ir elemento viršūnių koordinates.

2.14 lentelė. Ryšys tarp parametro reikšmės pokyčio ir elemento pokyčio pobūdžio
Table 2.14. The relation between change of parameter value and nature of element change

Elemento geometrijos tipas	Parametras	Elemento pokyčio pobūdis
Taškinis	Elemento geodezinės koordinatės x, y	Vietos pokytis
Linijinis	Elemento ilgis L	Konfigūracijos pokytis
Linijinis	Elemento pradžios taško geodezinės koordinatės x, y	Konfigūracijos arba vietos pokytis
Linijinis	Elemento pabaigos taško geodezinės koordinatės x, y	Konfigūracijos arba vietos pokytis
Linijinis	Elemento vidurio taško geodezinės koordinatės x, y	Konfigūracijos arba vietos pokytis
Linijinis	Elemento viršūnių taškų geodezinės koordinatės x, y	Konfigūracijos arba vietos pokytis
Plotinis	Elemento plotas S	Konfigūracijos pokytis
Plotinis	Elemento perimetro ilgis L	Konfigūracijos pokytis
Plotinis	Elemento pradžios taško geodezinės koordinatės x, y	Konfigūracijos arba vietos pokytis
Plotinis	Elemento pabaigos taško geodezinės koordinatės x, y	Konfigūracijos arba vietos pokytis
Plotinis	Elemento vidurio taško geodezinės koordinatės x, y	Konfigūracijos arba vietos pokytis
Plotinis	Elemento viršūnių taškų geodezinės koordinatės x, y	Konfigūracijos arba vietos pokytis

Išanalizavus erdvinių duomenų rinkinio GDR10LT elementų klases, vaizduojančias žemės dangos („Plotai“), kelių („Keliai“), hidrografijos („Hidro_l“) ir pastatų („Pastat“) objektus, dengiančius 25 km² teritoriją, nustatyta, kad priklausomai nuo elemento klasės elementą gali sudaryti nuo 2 iki 2920 viršūnių. Šio tyrimo rezultatai pateikti 2.15 lentelėje. Atsižvelgiant į šiuos rezultatus, nėra praktiška sudaryti identifikatorių, naudojant linijinio arba plotinio elemento viršūnių x, y koordinačių reikšmes. Todėl sudarant identifikatorių reikėtų naudoti kelias 2.14 lentelėje išvardintas parametrų reikšmes, išskyrus taškiniams elementams, taip sukuriant unikalias elemento savybes apibūdinančius identifikatorius, kurie keičiantis elemento konfigūracijai ar vietai, pakis.

2.15 lentelė. GDR10LT elementų klasių „Hidro_l“, „Keliai“, „Pastat“, „Plotai“ analizės rezultatai

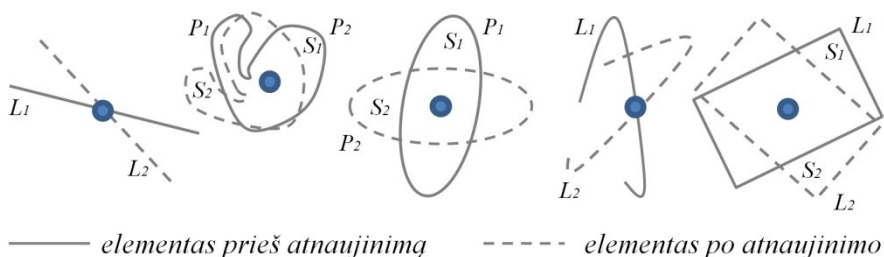
Table 2.15. The results of analysis of GDR10LT feature classes “Hidro_l”, “Keliai”, “Pastat”, “Plotai”

Elementų reikšmės		GDR10LT elementų klasės			
		„Hidro_l“	„Keliai“	„Pastat“	„Plotai“
kiekis		223	322	1448	896
Plotas, m ² (S)	Min.	–	–	0,34	0,047
	Maks.	–	–	2455,85	1842422,63
	Vidurkis	–	–	131,12	27901,78
perimetras/ ilgis, m (L)	Min.	1,84	3,25	3,2	3,49
	Maks.	1513,80	1684,13	357,73	17692,47
	Vidurkis	251,73	228,98	357,73	861,19
viršūnių kiekis	Min.	2	2	4	3
	Maks.	197	72	35	2920
	Vidurkis	16,30	9,86	6,48	124,44

Sudarant identifikatorių, priklausomai nuo elemento tipo, siūloma naudoti šias parametrų reikšmes:

- taškiniams elementams: elemento geodezinių koordinačių x, y reikšmės;
- linijiniams elementams: elemento ilgio, elemento centro geodezinių koordinačių x, y reikšmės;
- plotiniams elementams: elemento ploto, perimetro ir elemento centro geodezinių koordinačių x, y reikšmės.

Tačiau šis siūlomas metodas netinkamas sudaryti identifikatorių linijiniams ar plotiniams elementams, kai yra tikimybė, kad po pokyčio jie išlaikys tą pačią formą, o jų centro koordinatė nepakis (2.14 pav.).



2.14 pav. Elementų pokyčių pavyzdžiai, kai siūlomas identifikatorius nepakis
Fig. 2.14. The examples of elements changes when proposed identifier will not change

Kaip parodyta 2.15 paveiksle, elementų prieš ir po atnaujinimo CRC identifikatoriai sutaps, kadangi sutaps prieš ir po atnaujinimo:

- linijinių ar plotinių elementų centro taškų geodezinės koordinatės $X_1 = X_2$, $Y_1 = Y_2$;
- linijinių elementų ilgiai $L_1 = L_2$;
- plotinių elementų perimetrai $P_1 = P_2$;
- plotinių elementų plotai $S_1 = S_2$.

Esant tokiai tikimybei, papildomai CRC identifikatoriui sudaryti turi būti naudojamas:

- linijiniams elementams: elemento pradžios arba pabaigos taško geodezinių koordinatė x , y reikšmės;
- plotiniams elementams: elementą ribojančios linijos pradžios arba pabaigos taško, kurios visuomet sutampa, geodezinių koordinatė x , y reikšmės.

Erdvinių duomenų linijinių arba plotinių elementų centro geodezinėms koordinatėms nustatyti rekomenduojama naudoti formules, taikomas elemento centro arba masės centro koordinatė apskaičiavimui. Kadangi pasikeitus erdvinio duomenų elementų konfigūracijai ar vietai turi pasikeisti ir elemento centro koordinatės.

Taikytinos elemento centro arba masės centro koordinatė (C_x , C_y) apskaičiavimo formulės:

- plotiniams elementams (Bouek 1988):

$$C_x = \frac{1}{nS} \sum_{i=0}^{n-1} (x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i), \quad (2.15)$$

$$C_y = \frac{1}{nS} \sum_{i=0}^{n-1} (y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i), \quad (2.16)$$

čia S yra daugiakampio plotas apskaičiuojamas pagal formulę:

$$S = \frac{1}{2} \sum_{i=0}^{n-1} (x_i y_{i+1} - x_{i+1} y_i). \quad (2.17)$$

- linijiniams elementams (Protter, Morrey 1970):

$$C_x = \frac{x_1 + x_2 + \dots + x_n}{n}, \quad (2.18)$$

$$C_y = \frac{y_1 + y_2 + \dots + y_n}{n}, \quad (2.19)$$

čia n – linijinio elemento viršūnių skaičius.

- taškiniams elementams $C_x = x_i$, $C_y = y_i$.

Kadangi aukščiau minėtos reikšmės yra skaitinės, CRC identifikatoriui, skirtam elemento konfigūracijos ar vietos pokyčiui nustatyti, gali būti taikomos įvairios reikšmių kombinacijos: reikšmių sąjunga, skirtumas, sandauga ir pan. Sudarant CRC identifikatorių nerekomenduojama naudoti reikšmių sumos, kadangi taip gaunamos identifikatoriaus reikšmės prieš ir po atnaujinimo gali likti nepakitusios, nors realiai elementas bus pasikeitęs. Pavyzdžiui, nepakitusios konfigūracijos plotinis ar linijinis elementas perstumiamas į vietą, kur elemento centro X koordinatės reikšmė bus lygi ankstesnio elemento centro Y koordinatės reikšmei, o Y koordinatės reikšmė bus lygi ankstesnei X reikšmei. CRC identifikatoriui sukurti siūloma skaitines reikšmes interpretuoti, kaip tekstines ir iš jų sudaryti unikalią seką reikšmių, kuri identifikuotu kiekvieną elementą unikalčiai:

- taškiniams elementams [C_x 'AND C_y '];
- linijiniams elementams [C_x 'AND C_y 'AND L '];
- plotiniams elementams [C_x 'AND C_y 'AND P ' AND S '].

Prireikus nustatyti, ar pasikeitė erdvinių duomenų elementui priskirta savybės reikšmė, papildomai į CRC skaičiavimus būtina įtraukti savybės reikšmes.

Tačiau naudojant daug reikšmių identifikatoriui sudaryti, jo reikšmė bus labai ilga, todėl CRC identifikatoriaus reikšmei apskaičiuoti taikomas MD5 algoritmas, kuris reikšmes pakeičia į šešioliktainę sistemą.

Pavyzdžiui, skaičiuojant plotinių elementų CRC identifikatorių galima naudoti *Python* programavimo kalbą. MD5 algoritmas yra išreiškiamas taip:

```
def updateValue(lenval, areval, xval, yval):
    import hashlib
    strlenval = str(lenval)
    strareaval = str(areval)
    strx = str(xval)
    stry = str(yval)
    allstring = strlenval+strareaval+strx+stry
    m = hashlib.md5()
    m.update(allstring)
    return m.hexdigest()
```

Šiame scenarijuje kiekvieno elemento CRC identifikatoriui apskaičiuoti naudojama *updateValue* funkcija, kurioje:

- *lenval* – plotinio elemento perimetro reikšmė;
- *areaval* – plotinio elemento ploto reikšmė;
- *xval* – plotinio elemento centroidės koordinatė X reikšmė;
- *yval* – plotinio elemento centroidės koordinatės Y reikšmė.

Sukuriamas CRC identifikatorius pradiniam neatnaujintiems ir atnaujintiems erdvinio duomenų elementams. Tuo atveju, kai pradiniai erdviniai duomenų elementai jau turi unikalios identifikacijos priskirtus ir palaikomus pačiame duomenų rinkinyje, naujos ir senos versijų elementai turėtų būti apjungiami naudojant šiuos unikalios identifikacijos, o CRC identifikatorius būtų naudojamas tik elementų pokyčiams identifikuoti. Jeigu duomenų šaltinis nepalaiko unikalios identifikacijos, CRC identifikatorius turi būti taip pat naudojamas atnaujintų ir neatnaujintų elementų apjungimui. Tačiau jeigu duomenų šaltinis palaiko unikalios identifikacijos bei erdvinio duomenų elementų gyvavimo ciklo atributus, tuomet duomenų pokyčiai gali būti identifikuojami atsižvelgiant į elemento gyvavimo ciklo atributus. Priklausomai nuo to, ar erdvinio duomenų rinkinys palaiko unikalios identifikacijos ar ne, siūlomi du duomenų atnaujinimo aptikimo algoritmai. Abiem atvejais yra analizuojami neatnaujinti elementai, pirminiame etape saugomi tarpinių erdvinio duomenų bazėje ir atnaujinti elementai, kurie buvo perkelti į atnaujinimo šaltinių duomenų bazę. Tikslas yra nustatyti pokyčius ir atnaujinti tarpinių erdvinio duomenų bazėje saugomus erdvinio duomenis.

I duomenų atnaujinimo aptikimo algoritmas. Duomenų pokyčių aptikimo algoritmas, kai erdvinio duomenų šaltinyje palaikomas unikalios identifikacijos:

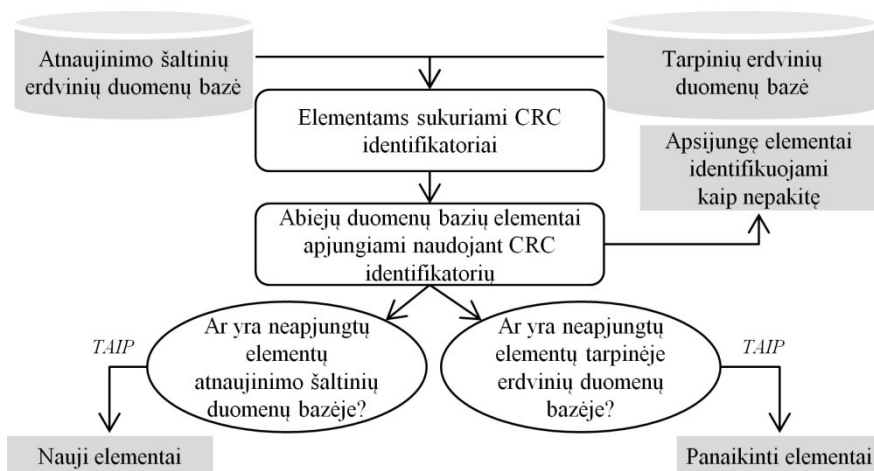
1. Kiekvienam elementui, esančiam neatnaujintame ir atnaujintame erdvinio duomenų rinkinyje, sukuriamas CRC identifikatorius.
2. Neatnaujinto ir atnaujinto rinkinio elementai sutapdinami (apjungiami) naudojant elementų unikalios identifikacijos.
3. Atliekama palyginimo operacija, ar apjungtų elementų CRC identifikatoriai sutampa. Jeigu sutampa, identifikuojami nepasikeitę elementai, kuriems įrašoma reikšmė „Nepakitęs“. Unikalus identifikatorius taip pat lieka nepakitęs. Jeigu CRC identifikatoriai nesutampa, identifikuojami pasikeitę elementai. Įrašoma reikšmė „Pasikeitęs“. Į tarpinių erdvinio duomenų bazę ir įkeliamas pasikeitęs elementas su savo atributais. Senos versijos elementas iš minėtos duomenų bazės yra ištrinamas. Elemento unikalios identifikacijos išlieka tas pats.
4. Atliekama tikrinimo operacija, ar liko neapjungtų elementų atnaujintame erdvinio duomenų rinkinyje. Jeigu liko, identifikuojami nauji elementai. Nauji elementai su savo unikalios identifikacijos

įkeliami į tarpinių erdvinių duomenų bazę ir jiems įrašoma reikšmė „Naujas“.

5. Atliekama tikrinimo operacija, ar liko neapjungtų elementų neatnaujintoje erdvinių duomenų rinkinyje. Jeigu liko, identifikuojami ištrinti elementai, įrašoma lauko reikšmė „Ištrintas“.

II duomenų atnaujinimų aptikimo algoritmas. Duomenų pokyčių aptikimo algoritmas, kai duomenų šaltinyje nepalaikomas unikalūs identifikatoriai (2.15 pav.):

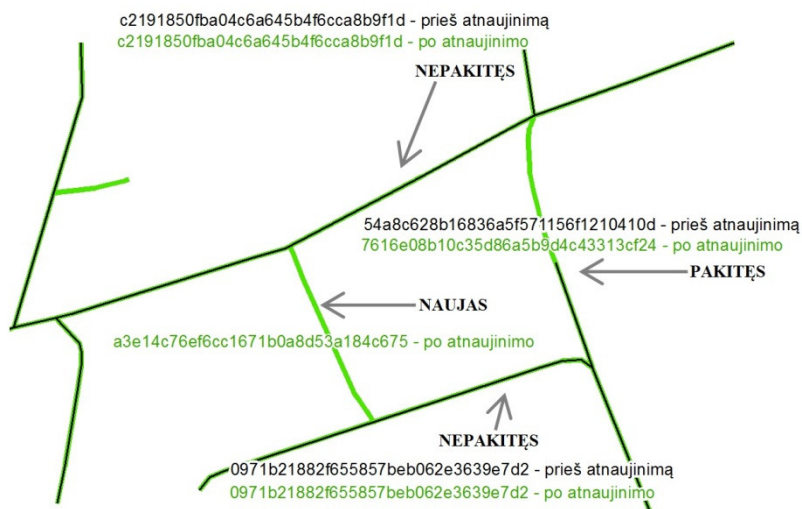
1. Kiekvienam naujos ir senos versijos elementui sukuriamas CRC identifikatorius.
2. Abiejų versijų objektai apjungiami naudojant CRC identifikatorius.
3. Atliekama palyginimo operacija, ar naujos versijos elemento CRC identifikatorius sutampa su senos versijos elemento CRC identifikatoriumi. Jeigu sutampa, identifikuojami nepasikeitę elementai ir tiems elementams įrašoma reikšmė „Nepakitęs“. Elemento identifikatorius lieka nepakitęs.
4. Jeigu CRC identifikatoriai nesutampa, atliekama tikrinimo operacija, ar liko neapjungtų elementų atnaujintame erdvinių duomenų rinkinyje. Jei liko, identifikuojami nauji elementai, įrašoma reikšmė „Naujas“. Taip pat atliekama tikrinimo operacija, ar liko neapjungtų elementų neatnaujintame erdvinių duomenų rinkinyje. Jeigu liko, identifikuojami ištrinti elementai, įrašoma reikšmė „Ištrintas“.



2.15 pav. Duomenų pokyčių aptikimo algoritmas nenaudojant unikalūs identifikatoriaus

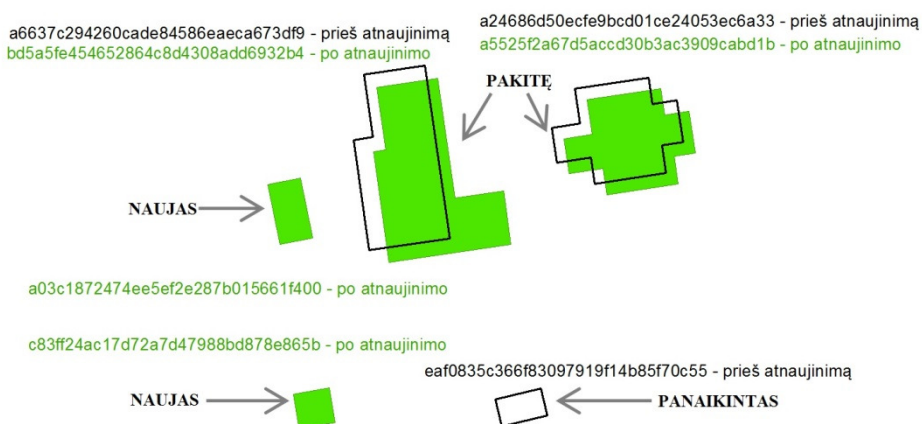
Fig. 2.15. The algorithm for detection of changes when unique identifier is not used

Siekiant įsitikinti šiame poskyryje siūlomo metodo patikimumu, GDR10LT erdvinį duomenų atnaujintų ir neatnaujintų elementų klasių „Keliai“, „Pastat“ ir „Plotai“ elementams testuojamoje teritorijoje priskirti CRC identifikatoriai. Apjungus atnaujintus ir neatnaujintus elementus, naudojant jiems priskirtus CRC identifikatorius, nustatyti elementai, kurių pasikeitė konfigūracija arba vieta (2.16–2.18 pav.).



2.16 pav. Identifikuotų elementų pokyčių pavyzdys („Keliai“)

Fig. 2.16. Example of detected changes in features („Keliai“)

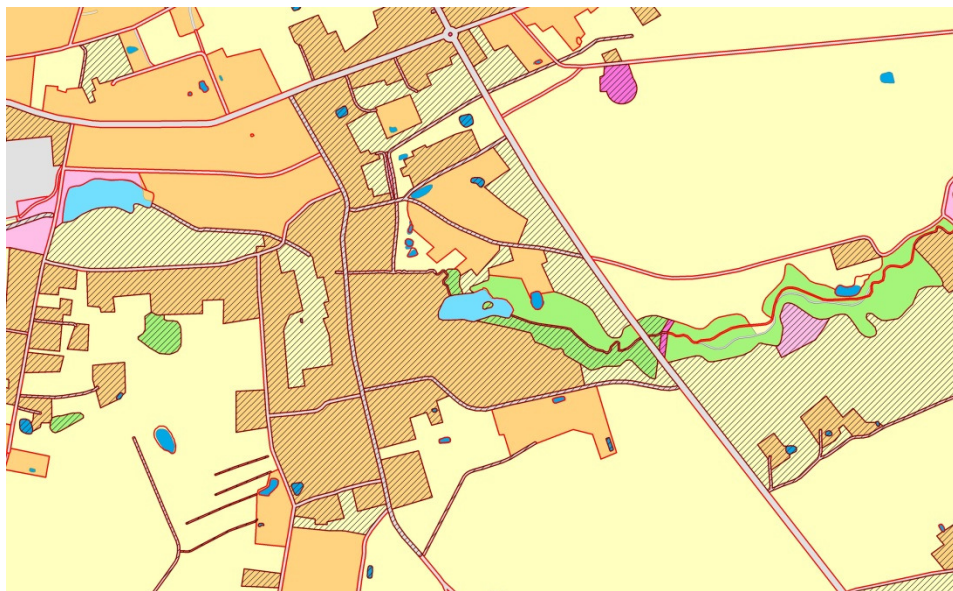


2.17 pav. Identifikuotų elementų pokyčių pavyzdys („Pastat“)

Fig. 2.17. Example of detected changes in features („Pastat“)

2.16 ir 2.17 paveiksluose juoda linija pavaizduoti elementai ar jų ribos GDR10LT elementų klasėse „Keliai“ ir „Pastat“ prieš atnaujinimą, žalia – po atnaujinimo. Kaip parodyta šiuose paveiksluose, pakitę elementai turės pasikeitusius CRC identifikatorius. Tuo tarpu naujų elementų CRC identifikatorius neturės atitikmens erdvinuose duomenyse prieš atnaujinimą, o panaikinti elementai – po atnaujinimo.

Taip pat CRC identifikatoriai buvo priskirti GDR10LT elementų klasės „Plotai“ elementams. Atsižvelgiant į juos 2.18 paveiksle raudona spalva pavaizduotos erdvių duomenų elementų ribos prieš atnaujinimą, o juodai brūkšniuotos teritorijos parodo nepakitusius elementus. Elementai, turintys skirtingas GKODAS reikšmes 2.18 paveiksle pavaizduoti skirtingomis spalvomis.



2.18 pav. Identifikuotų pokyčių pavyzdys („Plotai“)

Fig. 2.18. Example of detected changes (“Plotai”)

Atlikus tyrimus nustatyta, kad siūlomas erdvių duomenų pokyčių nustatymas cikliniu perteklinės informacijos tikrinimo metodu harmonizavimo procese leidžia greitai surasti pokyčius pradinių erdvių duomenų rinkiniuose prieš ir po atnaujinimo. Tai yra atlikus pradinių erdvių duomenų harmonizavimą, pakartotiniame harmonizavimo procese prieš atliekant transformavimą, naudojant CRC identifikatorių, galima rasti pakitusius elementus pradiname erdvių duomenų rinkinyje. Tačiau atlikus šių erdvių

duomenų pokyčių analizę, nustatyta, kad CRC identifikatorius identifikuoja ir pačius mažiausius pokyčius, kuriuos gali lemti x , y koordinačių reikšmių tūkstantosios dalies pasikeitimas. Todėl norint sumažinti kiekį pasikeitimų reikia papildomai įvertinti kitas elementų savybes, taip pat vadovautis 2.3.3 poskyryje aprašyta erdvinio duomenų elementų pasikeitimo reikšmingumo vertinimu.

Taip pat šis metodas gali būti taikomas ieškant pokyčių tarp skirtingų erdvinio duomenų, bet tik tuo atveju, kai yra naudojami identiški elementai abiejuose rinkiniuose.

Nustatyti erdvinio duomenų pasikeitimus tarp skirtingo mastelio erdvinio duomenų rinkinių taikant šį metodą praktiškai neįmanoma, kadangi erdviniai duomenys kitame mastelyje turės kitokią elementų konfigūraciją ar net gali būti vaizduojami kitokiu elemento tipu, todėl CRC identifikatoriai visuomet skirsis.

2.3.2. Erdvinių duomenų pokyčių įtakos transformavimo procesui vertinimo metodas

Pagrindinis uždavinys nustatant elementų pokyčius yra nuspręsti, kurie elementai pasikeitė lyginant juos su ankstesniąja erdvinio duomenų versija ir koks yra elementų pokyčių pobūdis, kurį iš principo galima įvertinti atsakius į kelis klausimus pateiktus 2.16 lentelėje.

2.16 lentelė. Galimi erdvinio duomenų pokyčių pobūdžiai

Table 2.16. Possible nature of changes of spatial data

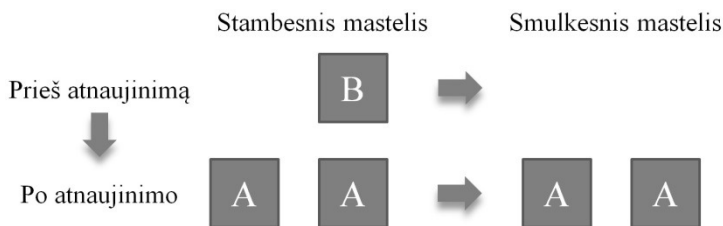
Klausimas	Pokyčio pobūdis	
	Pradinių erdvinio duomenų rinkinyje	Galutiniame erdvinio duomenų rinkinyje
Ar elementas yra naujas?	Naujas elementas	Naujas elementas
Ar pasikeitė elemento kokybinis požymis?	Pasikeitęs elemento kokybinis požymis	Naujas elementas arba pasikeitęs elemento kokybinis požymis
Ar elementas buvo panaikintas?	Panaikintas elementas	Panaikintas elementas arba pasikeitusi elemento konfigūracija (elementai neturėtų būti agreguojami)
Ar pasikeitė elemento konfigūracija?	Pasikeitusi elemento konfigūracija	Pasikeitusi elemento konfigūracija
Ar pasikeitė atstumas tarp dviejų elementų?	Pasikeitusi elemento konfigūracija	Pasikeitusi elemento konfigūracija (atlikti elemento agregavimą)

Pirmajame etape rekomenduojama atrinkti erdvinių duomenų pokyčius naudojant 2.3.1 poskyryje aprašytą metodą. Sekančiame etape turi būti analizuojamas rastų elemento pokyčių pobūdis, kuris leis jį įvertinti ir nustatyti įtaką harmonizavimo procesui ir parinkti transformavimo būdą. Vertinant erdvinių duomenų pokyčius būtina remtis pradinių ir galutinių duomenų rinkinių modelių analizės rezultatais, kurie turi būti vykdomi vadovaujantis 2.1 poskyryje aprašyta erdvinių duomenų harmonizavimo metodikos bendrąja dalimi.

Naujas elementas galutiniame erdvinių duomenų rinkinyje turės būti sukuriamas dviem atvejais, kai (2.19 pav.):

- nustatomas naujas elementas pradiniam erdvinių duomenų rinkinyje, kuriam yra priskirtos tokios savybių reikšmės, kurios yra vaizduojamos galutiniame erdvinių duomenų rinkinyje;
- esamas elementas, kuris yra vaizduojamas smulkesniame mastelyje, pradiniam erdvinių duomenų rinkinyje įgyja naują savybės reikšmę.

Šiais abiem atvejais iš pradinių erdvinių duomenų rinkinio turės būti atrinkti elementai, kuriuos transformavus, gautas rezultatas integruojamas į galutinį erdvinių duomenų rinkinį. Šiuo atveju jei vyksta harmonizavimas tarp skirtingo mastelio erdvinių duomenų rinkinių, turi būti papildomai taikomos generalizavimo procedūros.

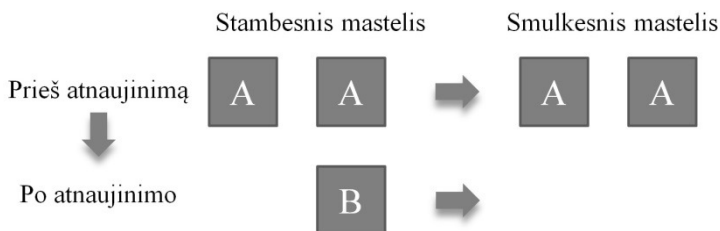


2.19 pav. Naujo elemento sukūrimo atvejai

Fig. 2.19. Cases of creation of new feature

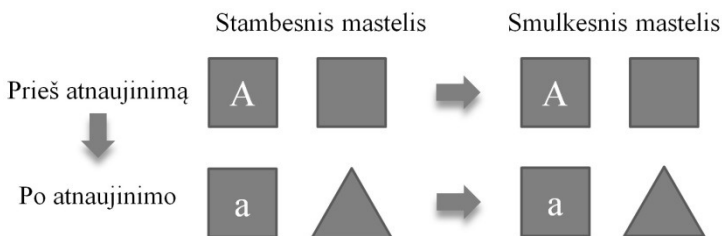
Elementas galutiniame erdvinių duomenų rinkinyje turės būti panaikintas priešingais atvejais, nei sukuriant naują elementą, t. y. (2.20 pav.):

- nustatomas panaikintas elementas pradiniam erdvinių duomenų rinkinyje, turėjęs tokią savybės reikšmę, kuri yra naudojama galutiniame erdvinių duomenų rinkinyje;
- pradiniam erdvinių duomenų rinkinyje esamas elementas įgyja naują savybės reikšmę, kuri nėra naudojama smulkesniame mastelyje.

**2.20 pav.** Elemento panaikinimo atvejai**Fig. 2.20.** Cases of deletion of feature

Jei pakitus elementui jis turi būti panaikintas, jokia generalizavimo procedūra nėra atliekama, išskyrus tais atvejais, kai elementas yra susijęs su kaimyniniais elementais.

Esamas elementas galutiniame erdvinio duomenų rinkinyje atnaujinamas, kai pasikeičia atitinkamo elemento savybės reikšmė, arba konfigūracija pradinio erdvinio duomenų rinkinyje (2.21 pav.). Pirmuoju atveju yra atnaujinami elemento požymiai, antruoju – konfigūracija.

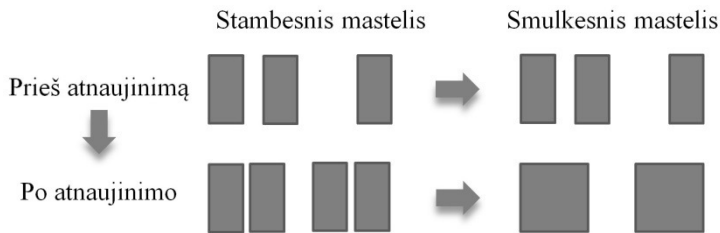
**2.21 pav.** Elemento atnaujinimo atvejai**Fig. 2.21.** Cases of update feature

Jei pasikeitė tik elemento savybių reikšmės, tuomet generalizavimo atlikti nereikia (atnaujinamas tik savybės reikšmė), o pasikeitus konfigūracijai būtina atlikti elemento supaprastinimą.

Nagrinėjant aukščiau minėtus atvejus, kai harmonizavimas vykdomas tarp skirtingo mastelio erdvinio duomenų rinkinių, papildomai reikia įvertinti atstumą iki kaimyninių elementų, turinčių tas pačias savybių reikšmes, ar jis yra didesnis ar mažesnis nei nustatytas minimalus leidžiamas atstumas:

1. Naujas elementas ar „priartėjęs“ esamas elementas prie kaimyninio elemento bus su juo agreguotas, tai yra elementas smulkesniame

mastelyje padidės (pavyzdžiui šalia buvusios užstatytos teritorijos atsirado naujas namų kvartalas) (2.22 pav.);



2.22 pav. Agreguoto elemento padidėjimo atvejai

Fig. 2.22. Cases of enlargement of an aggregated feature

2. Panaikinus elementą, arba „nutolus“ elementui, kuris anksčiau buvo agreguotas su kaimyniniu, jis turės būti eliminuotas iš agreguoto plotinio elemento galutiniame erdvinių duomenų rinkinyje, t. y. elementas sumažės (2.23 pav.).

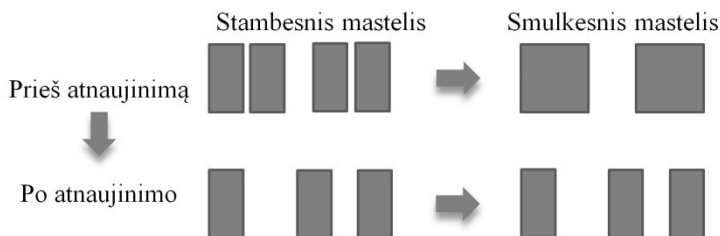


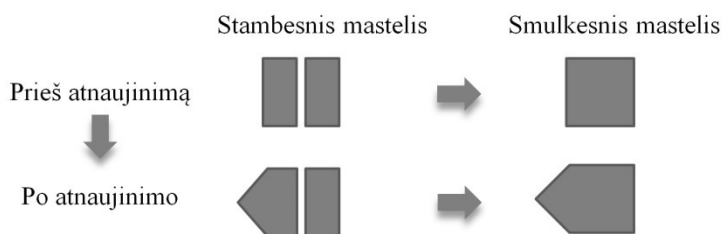
Fig 2.23. Agreguoto elemento sumažėjimo atvejai

Fig. 2.34. Cases of reduction of an aggregated feature

Pasikeitusios konfigūracijos elementas pradiniam erdvinių duomenų rinkinyje turės įtakos konfigūracijai to elemento, kuris buvo gautas agreguojant kaimyninius elementus (pavyzdžiui, iškirsta dalis vieno iš gretimų miškų) (2.24 pav.).

Nustačius erdvinių duomenų pokytį pirmuoju atveju, pavaizduotu 2.22 paveiksle, reikės atlikti elemento suprastinimo operaciją ir agregavimą su kaimyniniu elementu galutiniame erdvinių duomenų rinkinyje. Nors antruoju atveju (2.23 pav.) iš anksčiau agreguoto plotinio elemento galutiniame erdvinių duomenų rinkinyje reikėtų eliminuoti „netinkamą“ elementą, tačiau šiuo atveju vietoje eliminavimo funkcijos, kurią būtų sudėtinga įgyvendinti, siūloma iš

naudojant agregavimo funkciją likusiems elementams apjungti. Trečiajame variante (2.24 pav.), kaip ir pirmajame, reikia atlikti pasikeitusių elementų suprastinimą ir agregavimą su gretimais elementais.



2.24 pav. Agreguoto elemento formos pasikeitimo atvejai
Fig. 2.24. Cases of a change in the shape of aggregated feature

2.25 paveiksle pavaizduotas galimas ryšys tarp elemento pokyčio tipo ir erdvinio duomenų transformavimo būdo.

<i>sdata.v1</i>	Asfaltuotas kelias 	Asfaltuotas kelias		Asfaltuotas kelias
<i>sdata.v2</i>	Asfaltuotas kelias 	Asfaltuotas kelias 	Asfaltuotas kelias 	
<i>udata.v1</i>		Asfaltuotas kelias 		Asfaltuotas kelias
<i>udata.v2</i>	Asfaltuotas kelias 	Asfaltuotas kelias 	Asfaltuotas kelias 	
	<i>Sukurti naują elementą</i>	<i>Atnaujinti elemento konfigūraciją</i>	<i>Sukurti naują elementą</i>	<i>Panaikinti elementą</i>

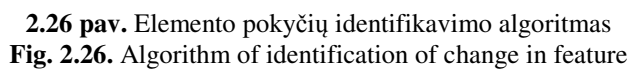
2.25 pav. Pavyzdys ryšio tarp elemento pokyčio tipo ir erdvinio duomenų transformavimo būdo

Fig. 2.25. Example of the relation between change of feature and way of transformation of spatial data

Erdvinio elemento savybės pokyčio, įtakojančio erdvinių duomenų harmonizavimą, identifikavimo algoritmas pateiktas 2.26 paveiksle.

Šiuo atveju siūloma atlikti veiksmus šia tvarka:

1. Sukuriama sąsaja tarp erdvinių duomenų elementų prieš ir po atnaujinimo. Šis veiksmas yra būtinas, jei elementai neturi unikalių identifikatorių. Šiame žingsnyje gali būti sukuriamas CRC identifikatorius.
2. Pokyčių identifikavimas:
 - 2.2. Susijusių erdvinių duomenų elementų (*sdata.r*) analizė ir erdvinių duomenų elementų, kurie pasikeitė po atnaujinimo, atranka. Erdvinių duomenų elementai su pakitusiomis savybių reikšmėmis patenka į grupę A1, o erdvinių duomenų elementai su pakitusia konfigūracija – į grupę A2.
 - 2.3. Erdvinių elementų (*sdata.v2*) po atnaujinimo, kurie neturi ryšio su erdviniais elementais prieš atnaujinimą, atranka. Šie elementai patenka į grupę B.
 - 2.4. Erdvinių elementų (*sdata.v2*) prieš atnaujinimą, kurie neturi ryšio su erdviniais elementais po atnaujinimo, atranka. Šie elementai patenka į grupę C.
3. Elementų esančių grupėse A1, A2, B ir C analizė ir vertinimas:
 - 3.1. Erdvinių duomenų elementų patekusių į grupes A1, B ir C analizė ir elementų su tinkamomis savybių reikšmėmis atranka. Kiti elementai yra pašalinami iš grupių.
 - 3.2. Erdvinių duomenų elementų, patekusių į grupę A2, analizė elementų ir elementų su reikšmingais konfigūracijos (geometrijos) pokyčiais atranka. Kiti objektai yra pašalinami iš grupės.
 - 3.3. Erdvinių duomenų elementų, patekusių į grupes B ir C, analizė ir elementų su tinkamomis geometrinėmis savybėmis atranka. Kiti elementai yra pašalinami iš grupių.
4. Ryšių tarp erdvinių duomenų elementų grupėse A1, A2, B ir C sukūrimas ir duomenų atnaujinimas (smulkesnio mastelio erdvinių duomenų elementai) (*Pokyčiai.udata.v1*).
5. Erdvinių duomenų elementų, esančių grupėse A1, A2, B ir C, kurie susiję su atnaujinamais duomenimis, analizė. Erdviniai duomenų elementai iš grupių:
 - 5.1. A1 ir B identifikuoja erdvinių duomenų elementų su atnaujintomis savybių reikšmėmis vietas.
 - 5.2. A2 identifikuoja vietas, kuriose atnaujintos erdvinių duomenų elementų konfigūracijos.



- 5.3. C identifikuoja pašalintų erdvinių duomenų elementų vietas.
6. Erdvinių duomenų elementų, esančių grupėse A1, A2, B ir C, kurie neturi ryšio su atnaujinamais duomenimis, analizė. Erdvinių duomenų elementai iš A1, A2, B grupių identifikuoja naujų erdvinių duomenų elementų sukūrimo vietas.

2.3.3. Erdvinių duomenų plotinių elementų pasikeitimo reikšmingumo vertinimo metodika

Identifikuojant erdvinių elementų pasikeitimą yra svarbu nustatyti reikšmingus pasikeitimus ir atmesti nereikšmingus. Pasikeitimai bus reikšmingi, kai:

1. įgyta nauja savybės reikšmė yra vaizduojama galutiniame erdvinių duomenų rinkinyje;
2. elemento konfigūracijos pasikeitimas bus matomas smulkesnio mastelio erdvinių duomenų rinkinyje.

Pasikeitimų reikšmingumo įvertinimas yra sudėtingas procesas, ypač norint įvertinti įtaką harmonizavimui tarp skirtingo mastelio erdvinių duomenų rinkinių. Šiuo atveju universalus metodo šiam procesui nėra, kadangi harmonizuojant erdvinius duomenis yra taikomas generalizavimo procesas, kuriam parenkami algoritmai priklauso nuo pradinių ir galutinių erdvinių duomenų modelių.

Bendruoju atveju vertinant elementų pasikeitimo reikšmingumą turi būti tiriamas analizuojamas pasikeitęs pavienis elementas ir kaip jo pasikeitimas įtakos aplinkinius elementus pagal 2.3.2 poskyryje išskirtus pasikeitimų pobūdžius. Vertinimas turi būti atliekamas tokiose elementų pokyčių grupėse tokia eilės tvarka:

1. elementai turintys pakitusius jiems priskirtų savybių reikšmes;
2. nauji elementai;
3. panaikinti elementai;
4. kaimyniniai elementai, kurie turi būti agreguojami;
5. pasikeitusios konfigūracijos elementai.

Elemento požymio atnaujinimas. Vertinant pradinio erdvinių duomenų rinkinio savybės reikšmės pasikeitimo reikšmingumą, užtenka žinoti, kokios savybių reikšmės yra svarbios, tai yra priskiriamos galutinio erdvinių duomenų rinkinio elementams ir lyginant elementą prieš ir po atnaujinimo pradiname erdvinių duomenų rinkinyje išskiriami elementai, kurie įgijo naujas svarbias savybių reikšmes. Pavyzdžiui, patikslintas ežero pavadinimas.

Elemento sukūrimas. Naujo elemento atsiradimas bus reikšmingas panašiais atvejais kaip ir elemento savybės reikšmės pasikeitimo atveju, tai yra, jei naujas elementas turės tinkamas savybių reikšmes.

Elemento panaikinimas. Elemento panaikinimas pradiniam erdvinį duomenų rinkinyje bus reikšmingas, jei elementas anksčiau buvo saugomas galutiniame erdvinį duomenų rinkinyje.

Agreguojamas elementas. Pakitęs elementas agregavimo operacijos atžvilgiu bus reikšmingas, jei jis yra „priartėjęs“ arčiau nei leistinas minimumas tarp kaimyninių elementų, turinčių tokias pačias savybių reikšmes. Tokių pokyčių paieška gali būti lengvai įgyvendinama naudojant erdvinės analizės užklausą, kurios metu ieškomos sankirtos tarp sukurto buferio aplink pakitusios konfigūracijos elementą ir kaimyninio elemento. Sukuriama elemento plotis turi būti lygus minimaliam leidžiamam atstumui tarp kaimyninių elementų.

Priartėjusių elementų paieškos algoritmas, įgyvendinamas naudojant *ArcGIS ModelBuilder* aplinką, pateiktas 2.17 lentelėje.

2.17 lentelė. Priartėjusių elementų paieškos algoritmas

Table 2.17. Algorithm of identification approached of features

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Atrenkami pakitę elementai (naudojamas CRC identifikatorius).	<i>Select by Attribute</i>
2.	Pakitę elementai perkeliama į naują erdvinį duomenų elementų klasę.	<i>Feature Class to Feature Class</i>
3.	Aplink pakitusius elementus sukuriamas buferis.	<i>Create Buffer</i>
4.	Atrenkami antrame žingsnyje sukurti buferiai, kurie kertasi su nepakitusiais elementais.	<i>Select by Location</i>
5.	Atrenkami antrame žingsnyje sukurti buferiai, kurie kertasi vienas su kitu.	Nustatyti galima sukuriant modelį naudojant <i>ArcGIS</i> topologines taisykles (<i>Topology</i>).
6.	Naudojant ketvirtame ir penktame žingsniuose atrinktus buferius atrenkami antrame žingsnyje atrinkti pakitę elementai.	<i>Select by Location</i>

Lentelėje 2.17 atrinkti elementai bus svarbūs erdvinį duomenų harmonizavimo procesui.

Elemento konfigūracijos keitimas. Elemento konfigūracijos pasikeitimo reikšmingumo nustatymas yra sudėtingesnis, nei anksčiau aprašyti atvejais. Plotinio elemento konfigūracijos pokyčio reikšmingumo vertinimui siūloma:

1. taikyti elementų sankirtos ir erdvinės analizės užklausas,
2. įvertinti elemento pokyčio ploto ir perimetro santykį (elemento formą).

Elemento konfigūracijos pasikeitimo reikšmingumas turi būti nustatomas lyginant pasikeitimo dydį su nustatytu leidžiamu minimaliu pasikeitimu, kuris pirmiausiai turėtų priklausyti nuo galutinio erdvinių duomenų rinkinio mastelio ir erdvinių duomenų specifiškumo. Papildomai turi būti įvertinamas tikėtinas elemento dydis prieš ir po generalizavimo.

Prieš pradedant analizuoti elemento konfigūracijos pokyčius abiem atvejais iškart atmetami pakitusios konfigūracijos elementai, kurie pagal jiems priskirtas savybių reikšmes nėra vaizduojami galutiniame erdvinių duomenų rinkinyje.

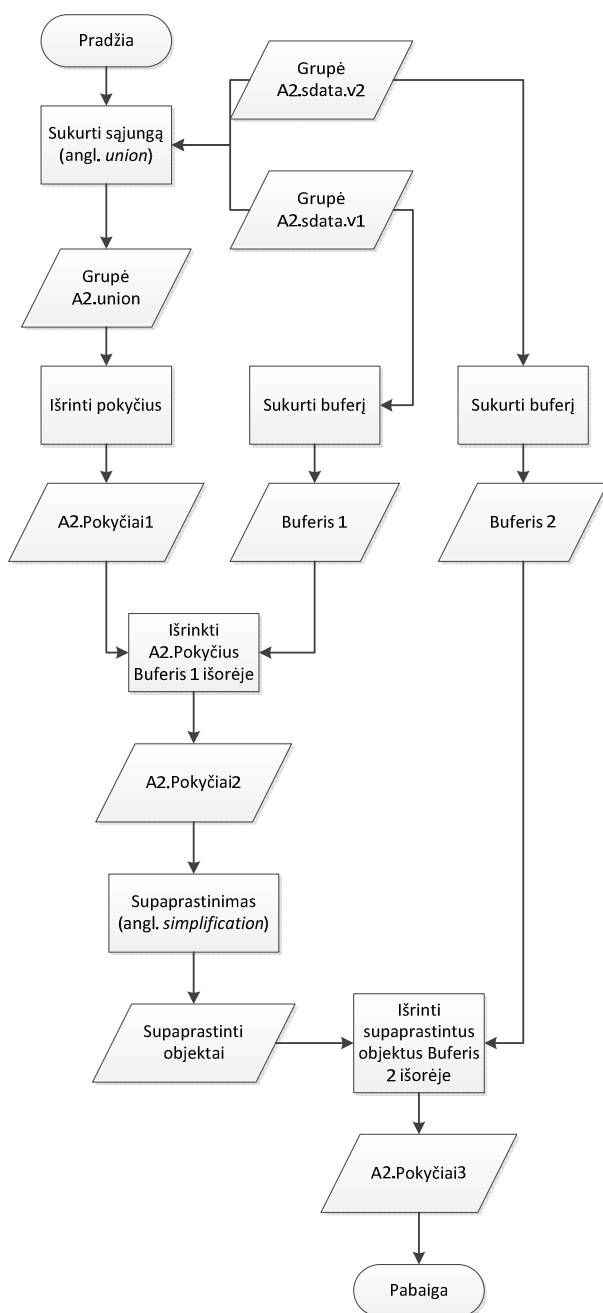
Atliekant analizę pirmuoju atveju reikia taikyti algoritmą, skirtą palyginti pradinių erdvinių duomenų rinkinio elementams prieš ir po atnaujinimo, rasti elemento konfigūracijos pasikeitimus, juos įvertinti ir atrinkti reikšmingus. Naudojant GIS technologijas šį uždavinį išspręsti galima taikant erdvinę analizę, kurioje veiksmai turi būti atliekami tokia tvarka (2.27 pav.):

1. susiejami erdviniai duomenys prieš (*A2.sdata.v1*) ir po atnaujinimo (*A2.sdata.v2*). Atrenkamos pakitusio dalys (*A2.changes1*). Susiejimas gali būti vykdomas taikant 2.3.1 poskyryje pateiktą erdvinių duomenų pokyčių nustatymo cikliniu perteklinės informacijos tikrinimo metodu.
2. Buferio pločio reikšmės nustatymas. Šio parametro dydis priklauso nuo erdvinių duomenų kartografinio mastelio ir reikalavimų erdvinių duomenų elementų minimaliems dydžiams. Klasikinėje kartografijoje, atsižvelgiant į kartografinį mastelį žemėlapio tikslumas milimetrais nustatomas:

$$t = 0,1 \cdot M, mm, \quad (2.20)$$

čia M – kartografinio mastelio vardiklis.

3. Buferinių zonų sukūrimas, pagal nustatytą parametą, aplink pirminius neatnaujintus erdvinis duomenis (*Buffer1*).
4. Elementų, esančių už buferio ribos, atrinkimas (*A2.changes2*).
5. Atrinktų elementų prastinimas (*SimplifyObject*).
6. Buferinių zonų sukūrimas, pagal nustatytą parametą, aplink atnaujintų erdvinių duomenų elementus (*Buffer2*).
7. Atrenkami elementai po suprastinimo, esantys už buferinės zonos ribų (*A2.changes3*). Atrinkti pasikeitimai bus reikšmingi.



2.27 pav. Reikšmingų elementų pokyčių paieškos algoritmas
Fig. 2.27. The algorithm of identification of significant changes in features

Atliekant analizę antruoju atveju naudojamas ankstesnio erdvinės analizės uždavinio pirmajame žingsnyje gauto rezultato elementas (pokytis) *A2.changes1*. Vertinant pokyčio formą, darome prielaidą, kad kuo elementas yra labiau apskritas tuo didesnė tikimybė, kad jis turės būti generalizuojamas. Elemento ar jo pokyčio konfigūracija įvertinama nustatant siaurumo mastą (angl. *Thickness Ratio*):

$$T = 4\pi \frac{S}{P^2}, \quad (2.21)$$

čia S – plotinio elemento plotas, P – plotinio elemento perimetras.

Kuo T yra arčiau 1, tuo plotinis elementas yra panašesnis į apskritimą. Tačiau šis būdas yra netinkamas įvertinti plotinių elementų formos, kurio viduje yra „skylių“, kadangi šiuo atveju padidėja perimetro ilgis. Todėl formulėje reikia taikyti plotinio elemento išorinį perimetrą arba prieš atliekant vertinimą eliminuoti „skyles“ elemento viduje.

Taikant šiame poskyryje siūlomus erdvinių duomenų pokyčių įvertinimo būdus, galima harmonizavimo procese aktyvuoti erdvinių duomenų generalizavimą, tik tuo atveju, kai taikant šį metodą bus nustatyti „reikšmingi“ pokyčiai.

Taikant erdvinių duomenų elementų sankirtą ir erdvinės analizės užklausas pokyčiams nustatyti, buvo lyginti GDR10LT erdvinių duomenų elementų klasės „Plotai“ elementai prieš ir po atnaujinimo. Iš viso prieš atnaujinimą buvo 3927 elementai, po atnaujinimo – 3955. Atlikus jų sąjungą (angl. *union*) buvo gautos 4916 elementų kombinacijos, turinčios kelių elementų savybių reikšmes. Elementai paveldėję ankstesnio elemento savybes parodo nepakitusias elementų vietas. Priešingu atveju – identifikuojamas pokytis. CRC identifikatorių analizė parodė pokytį 1648 elementuose. Atlikus jų supaprastinimą, atlikta erdvinių duomenų atranka naudojant sukurtus buferius, taikant pasirinktą 5 metrų atstumą. Toks atstumas pasirinktas, kadangi buvo vertinama, ar elemento konfigūracijos pokytis turės įtakos M1:50 000 erdvinių duomenų rinkiniui. Pašalinus pakitusias elementų dalis, patenkančias į buferį liko 952 elementų pokyčių, kurie gali turėti įtaką erdvinių duomenų M1:50 000 pokyčiams. Tai yra buvo identifikuota mažiau nei 25% elementai, kuriuos transformuojant reikėtų atlikti supaprastinimą.

2.4. Harmonizuojamų erdvinių duomenų modeliavimo tobulinimas

Erdvinių duomenų modeliavimas detaliam aptartas *INSPIRE Generic Conceptual Model* dokumente (toliau – INSPIRE koncepcinio modelio aprašas). Šis dokumentas numato detalias gaires aprašant INSPIRE temų erdvinių duomenų specifikacijas, tai yra galutinių erdvinių duomenų rinkinių modelius. Nors harmonizavimo procesas leidžia kaupti pradinis erdvinius duomenis, neatsižvelgiant į galutinių erdvinių duomenų modelius ir juose naudojamus klasifikatorius, atsižvelgiant į INSPIRE koncepcinio modelio aprašą ir į 2.3 poskyryje siūlomus erdvinių duomenų pokyčių nustatymo ir įvertinimo metodus, rekomenduojama modeliuojant erdvinius duomenis, kurie bus naudojami harmonizavimo procese, numatyti elementams šias savybes:

1. Elementų unikalūs identifikatoriai, kurie turi būti palaikomi visą jų gyvavimo ciklą.
2. Elementų gyvavimo ciklo savybės (elemento sukūrimo, pakeitimo ir panaikinimo data).
3. Sąsajas su galutinio erdvinių duomenų rinkinio tapačiais elementais.

INSPIRE apimtyje elementų unikalūs identifikatoriai yra paremtas INSPIRE direktyvos 8 straipsnio 2 dalies (a) punktu, kuris numato, kad erdvinių duomenų elementai turi būti unikalūs identifikuojami. Kuriant unikalūs identifikatoriai reikėtų atsižvelgti į jų (INSPIRE Generic... 2014):

1. Unikumą. Jokie du elementai erdvinių duomenų rinkinyje negali turėti tokių pat elementų identifikatorių. Elementui nustojus egzistuoti identifikatoriai negali būti iš naujo panaudojami.
2. Stabilumą. Sukurti elementų identifikatoriai turi išlikti tokie pat visą elemento egzistavimo laiką. Specifikuojant erdvinius duomenis taip pat turi būti aprašyta, kurie pasikeitimai (pavyzdžiui, savybės reikšmės pakeitimas, apjungimas su kitu erdviniu duomenų elementu) gali pakeisti elemento identifikatorių.
3. Suradimą (atsekimą). Elementai turi būti randami naudojant jų unikalūs identifikatoriai. Tai yra identifikatoriai turi suteikti informaciją apie pradinį erdvinių duomenų rinkinį. Tokiu būdu, prireikus galima būtų nesudėtingai rasti prieigą prie erdvinių duomenų šaltinio, skirtingų erdvinių duomenų rinkinių harmonizavimą.

2011 metais Lietuvos erdvinės informacijos išvystymo projekto metu remiantis INSPIRE principais buvo sukurtas Vieningas georeferencinių erdvinių duomenų modelis (toliau – VGDM), Valstybinio lygmens georeferencinių duomenų vieningam aprašymui. Šiame modelyje siūlomos elemento gyvavimo ciklo savybės pateiktos 2.18 lentelėje.

2.18 lentelė. Vieningo georeferencinio duomenų modelio erdvinių duomenų elementų gyvavimo ciklo savybės

Table 2.18. The attributes of life-cycle of elements of spatial data model for unify model of (geo)referential data model

Savybės kodas	Savybės pavadinimas	Savybės aprašymas
top-id	Unikalus identifikatorius	Unikali nuoroda, kuri neapima jokios informacijos apie elementą ir išlieka ta pati visame elemento egzistavimo cikle.
Reikšmių formatas	Tekstas	
Reikšmių sritis	Netaikoma	
objektoPradžiosLaikas	Elemento gyvavimo pradžios laikas	Data nurodanti, kada elementas yra sukurtas erdvinių duomenų rinkinyje (tai atitinka elemento egzistavimo ciklo pradžią).
Reikšmių formatas	Data	
Reikšmių sritis	Netaikoma	
objektoPabaigosLaikas	Elemento gyvavimo pabaigos laikas	Data nurodanti elemento egzistavimo ciklo erdvinių duomenų rinkinyje pabaigą.
Reikšmių formatas	Data	
Reikšmių sritis	Netaikoma	
versijosNumeris	Versijos numeris	Numeris nurodantis elemento versiją. Elemento egzistavimo ciklo pradžioje suteikiamas versijos numeris lygus 1.
Reikšmių formatas	Sveikas skaičius	
Reikšmių sritis	1 – 32767	
versijosData	Versijos data	Data nurodanti, kada buvo sukurta nauja elemento versija. Elemento egzistavimo ciklo pradžioje versijos data atitinka elemento sukūrimo erdvinių duomenų rinkinyje pradžią.
Reikšmių formatas	Data	
Reikšmių sritis	Netaikoma	

Šios siūlomos savybės atitinka INSPIRE koncepcinį duomenų modelį (INSPIRE Generic... 2014). Tačiau versijos numeris yra perteklinis, ypač esant dideliui kiekiui to paties pradinių erdvinių duomenų elemento pasikeitimų skaičiui. Taip pat vietoje versijos datos, rekomenduotina kaupti elemento pasikeitimo datą, kuri iš principo ir būtų versijos data. Modeliuojant harmonizuojamus erdvinius duomenis reikėtų numatyti savybes, kuriose būtų kaupiama informacija apie elementų pasikeitimo pobūdį ir priežastis. Tai dar

labiau pagreintų erdvinių duomenų elementų pokyčių identifikavimo ir įvertinimo procesą.

2.5. Antrojo skyriaus išvados

1. Parengta patobulinta erdvinių duomenų harmonizavimo metodika, praplečiant ir detalizuojant esamas bendrąsias metodikas, skirtas erdvinių duomenų harmonizavimui erdvinės informacijos infrastruktūroje. Joje numatyta papildoma erdvinių duomenų pokyčių nustatymo ir vertinimo procedūra, detalizuotas plotinių elementų konfigūracijos pokyčio įvertinimas. Pasiūlytas erdvinių duomenų modeliavimo patobulinimas, skirtas kaupti informaciją apie erdvinių duomenų elementų pokyčius ir jų pobūdį.
2. Užtikrinant harmonizavimo procesą, būtina papildomai įdiegti atnaujinimo šaltinių ir galutinių erdvinių duomenų patikros procedūras. Pradinių duomenų patikra užtikrins sukurto transformavimo proceso veikimą, kadangi pakitusi pradinių erdvinių duomenų struktūra ar pačių erdvinių duomenų nekorektiškumas, gali įtakoti transformavimo procesą: jis gali neveikti arba gali būti gaunamas netinkamas rezultatas. Galutinių erdvinių duomenų patikra, leis įvertinti atlikto transformavimo proceso teisingumą. Neigiamas patikros rezultatas reikštų, kad pradiniuose duomenyse gali būti pasikeitimų, kurie turėjo įtakos transformavimo rezultatui.
3. Erdvinių duomenų generalizavimas yra sudėtingas daugiakopis uždavinys, todėl transformavimo metu vykdant erdvinių duomenų generalizavimą, rekomenduojama šį procesą atskirti nuo kitų erdvinių duomenų transformavimo procesų. Taip būtų sudaroma galimybė pasikeitus generalizavimo parametrams, tikslinti ir testuoti tik erdvinių duomenų generalizavimo modelius.
4. Generalizavimo procese erdvinių duomenų plotinių elementų eliminavimui siūlomas algoritmas, kuriame pagrindinės transformavimo operacijos taikomos ne patiems plotiniams elementams, o jas ribojantiems linijiniams elementams, papildomai naudojant kitose erdvinių duomenų elementų klasėse esančius linijinius elementus eliminuojamų plotinių elementų padalinimui. Šis algoritmas ne tik eliminuos plotinius elementus, jų plotą prijungiant prie kaimyninio elementų, bet ir pastarųjų kraštinės po generalizavimo sutaps su egzistuojančiu linijiniu elementu, naudotu eliminuojamo elemento padalinimui.

5. Atlikus eksperimentinius tyrimus, nustatyta, kad erdviniams duomenims taikant Wang algoritmą, bendra suprastintų elementų vidutinė kokybės išraiška yra apie 0,9, praplėstą Wang algoritmą – 0,7, o taikant Douglas-Peucker algoritmą 0,56, kai tolerancijos reikšmė yra 20 m, 0,59, kai 10 m ir 0,63, kai 5 m. Atsižvelgus į tyrimų rezultatus, siūloma pirmuosius algoritmus taikyti gamtinių objektų generalizavimo modeliuose, o Douglas-Peucker – antropogeninių objektų suprastinime, atitinkamai parinkus tinkamą tolerancijos reikšmę, optimaliai išlaikant suprastinamų elementų kokybę.
6. Disertacijoje siūlomas plotinių elementų suprastinimo procesas, priklausomai nuo poreikio, gali suprastinti, tiek plotinius elementus, tiek plotinius elementus ribojančias linijas, taikant skirtingus suprastinimo algoritmus ir juose taikomas parametrų reikšmes. Atlikus suprastinimą proceso metu gauti persidengimai tarp elementų yra eliminuojami, juos prijungiant prioriteto tvarka prie kaimyninių elementų.
7. Siūlomas erdvinių duomenų kaimyninių elementų sutapdinimo algoritmas, kuris vietoje kaimyninių elementų kraštinių sutapdinimo, vieno elemento kraštinę pakeičia kaimyninio elemento kraštine. Taip yra užtikrinamas idealus kraštinių sutapimas.
8. Pasiūlyti sprendimai erdvinių duomenų elementų eliminavimui ir apjungimui, linijinių elementų atšakų eliminavimui ir elementų sutapdinimui, kurie kuriant erdvinių duomenų harmonizavimo technologijas gali būti taikomi transformavimo metu atliekant erdvinių duomenų generalizavimą kartu su kitais generalizavimo algoritmais.
9. Parengtas erdvinių duomenų pokyčių matavimo ir reikšmingumo įvertinimas, priklauso nuo jam priskirtos pokyčio klasės ir erdviniams duomenims taikomo kartografinio mastelio.
10. Pasiūlyti patobulinti erdvinių duomenų pokyčių nustatymo ir įvertinimo metodai, kurie įtraukti į siūlomą erdvinių duomenų harmonizavimo metodiką. Šie metodai leidžia įgyvendinti dalinį erdvinių duomenų generalizavimą.

Erdvinių duomenų harmonizavimo technologijų kūrimas

Skyriuje aprašomas erdvinių duomenų harmonizavimo technologijos sukūrimas, taikant disertacijos metu paruoštą patobulintą erdvinių duomenų harmonizavimo metodiką. Pateikiami realizuotų modelių bendrieji aprašymai ir atliktų modelių testavimo rezultatai.

Skyriaus tematika atlikti tyrimai paskelbti autorės straipsniuose (Papšienė *et al.* 2014, Papšienė 2013a).

3.1. Georeferencinių erdvinių duomenų harmonizavimo technologijos kūrimas

Remiantis antrajame skyriuje pateikta siūloma erdvinių duomenų harmonizavimo metodika sukurta technologija, skirta GDR10LT ir GDR50LT erdvinių duomenų elementų klasių, vaizduojančių kelius, upes ir žemės dangą harmonizavimui, kurią galima naudoti EII aplinkoje. Kai kurie aprašyti sprendimai yra skirti visoms erdvinių duomenų rinkinio GDR10LT elementų klasėms.

Vadovaujantis 2.1 poskyryje aprašyta metodika, pirmajame etape atlikta detali pradinių erdvinio duomenų rinkinio (GDR10LT) ir galutinio erdvinio duomenų rinkinio (GDR50LT) elementų klasių „Keliai“, „Hidro_l“ ir „Plotai“ analizė. A ir B prieduose pateikti erdvinio duomenų rinkinių GDR10LT ir GDR50LT duomenų modelių ištraukos. Apibendrinti analizės rezultatai pateikti 3.1 lentelėje.

3.1 lentelė. GDR10LT ir GDR50LT elementų klasių „Keliai“, „Hidro_l“, „Plotai“ analizės rezultatai

Table 3.1. The results of analysis of the feature classes “Keliai”, “Hidro_l”, “Plotai” in GDR10LT and GDR50LT

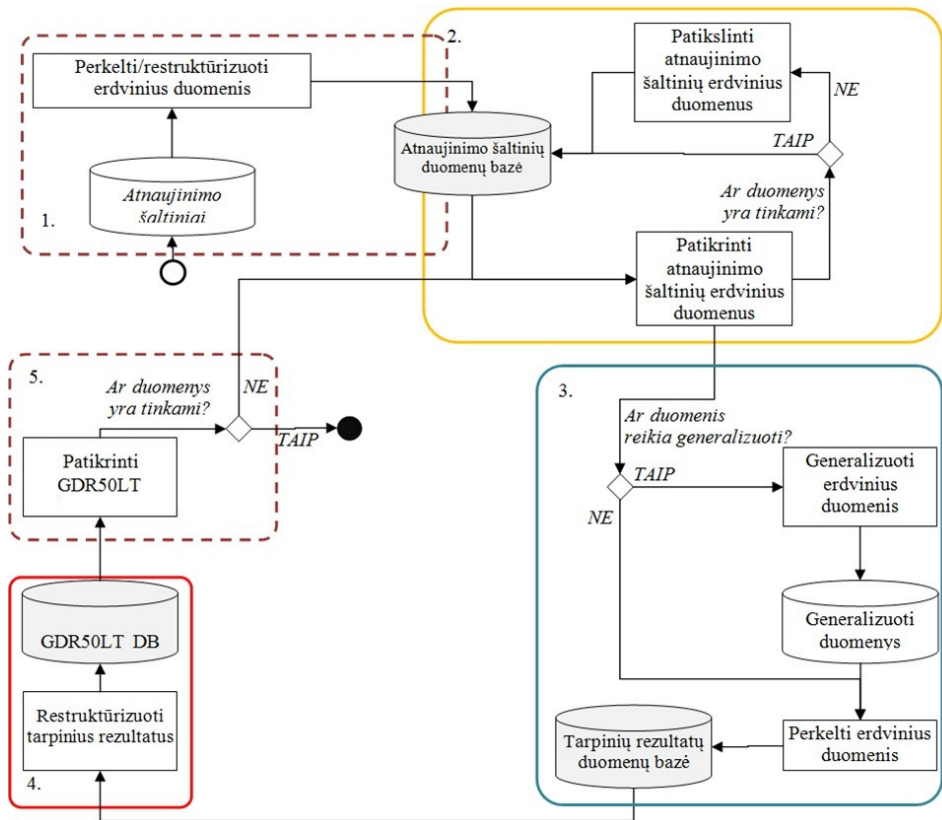
Eil. Nr.	Analizės uždaviniai	GDR10LT	GDR50LT	Išvada
1.	Erdvinių duomenų rinkinio analizė			
1.1.	Prieiga	Tiesioginė prieiga prie duomenų bazės (saugoma Lietuvos EII)		Erdvinių duomenų transformaciją galima vykdyti EII aplinkoje neinicijuojant pradinių erdvinį duomenų iš išorės gavimą.
1.2.	Koordinačių sistema	LKS-94		Sutampa
1.3.	Matavimo vienetų sistema	Metrinė		Sutampa
1.4.	Formatas	File Geodatabase		Sutampa
2.	Erdvinių duomenų rinkinių modelių analizė			
2.1	Terminologija	Taikoma vienoda terminologija		
2.2.	Erdvinių duomenų elementų klasių analizė	„Keliai“: važiuojamosios kelio dalies ašinės linijos. „Hidro_l“: upių, kanalų ir melioracijos griovių ašinės linijos. „Plotai“: žemės naudmenos ir vandens telkiniai.	„Keliai“: kelio ašinės linijos. „Hidro_l“: upių, kanalų ir melioracijos griovių ašinės linijos. „Plotai“: žemės naudmenos ir vandens telkiniai.	GDR10LT erdvinį duomenų elementų klasėje „Keliai“ yra kaupiami elementai, vaizduojantys važiuojamosios kelio dalies ašines linijas, tuo tarpu GDR50LT – kelio ašinės linijos. Todėl reikalinga atlikti lygiagrečių elementų sutraukimą į vieną linijinį elementą.

3.1 lentelės pabaiga

Eil. Nr.	Analizės uždaviniai	GDR10LT	GDR50LT	Išvada
2.3.	Klasifikatoriai	GDR10LT erdvinių duomenų elementų klasės „Keliai“, „Hidro_1“, „Plotai“ naudoja tuos pačius klasifikatorius.		
2.4.	Tikslumas	1:10 000	1:50 000	Transformavimo procese būtina naudoti erdvinių duomenų generalizavimo operacijas.
2.5.	Detalumas	Neapibrėžtas	Neapibrėžtas	
3.	Elementams priskirtų savybių reikšmių analizė:			
	GDR10LT erdvinių duomenų elementų klasės „Keliai“, „Hidro_1“, „Plotai“ naudoja tuos pačius klasifikatorius. Klasifikatoriaus „Plotai“: GKODAS reikšmės „gt0“, „hd0“, neturi jokio atitikmens GDR50LT klasifikatoriuje „Plotai“: GKODAS (1:0), todėl šie elementai, turi būti eliminuojami. GDR10LT „Plotai“: GKODAS reikšmės „va1“, „va11“ turės atitikmenį GDR50LT „Plotai“: GKODAS „va1“ (N:1).			
4.	Topologiniai ryšiai			
4.1.	Topologinės taisyklės	„Plotai“: elementai negali dengti vienas kito, tarp elementų negali būti skylių. „Keliai“, „Hidro_1“: elementai negali kirsti vienas kito, negali būti psiaudo taškų.	„Plotai“: elementai negali dengti vienas kito, tarp elementų negali būti skylių. „Keliai“, „Hidro_1“: elementai negali kirsti vienas kito, negali būti psiaudo taškų.	Atliekant erdvinių duomenų transformavimą, būtina atsižvelgti į GDR50LT erdviniams duomenims taikomas topologines taisykles.

Atlikus analizę nustatyta, kad abu erdvinių duomenų rinkiniai yra saugomi Lietuvos erdvinės informacijos infrastruktūroje, todėl nereikės naudoti LEI portalo elektroninių paslaugų pradiniais erdviniams duomenims pasiekti ar perduoti į galutinių erdvinių duomenų rinkinį. Šiuo atveju gali būti naudojama tiesioginė prieiga prie GDR10LT ir GDR50LT duomenų bazių. Taip pat nustatyta, kad sutampa erdvinių duomenų koordinacijų sistema, matavimo vienetai ir duomenų saugojimo formatas, todėl kuriant transformavimo procesą nereikia numatyti jų keitimo. Kadangi skiriasi GDR10LT ir GDR50LT kartografiniai masteliai, transformuojant duomenis būtina atlikti erdvinių duomenų generalizavimą.

Atsižvelgiant į analizės rezultatus ir disertacijoje pasiūlytą metodiką, transformavimo procese numatyti penki pagrindiniai etapai: atnaujinimo šaltinių paruošimo, atnaujinimo šaltinių patikros, erdviųjų duomenų generalizavimo, erdviųjų duomenų restruktūrizavimo ir automatizuotai atnaujinto GDR50LT patikros. Principinė transformavimo proceso schema, apimanti visus aukščiau paminėtus penkis etapus, pateikta 3.1 paveiksle.



3.1 pav. Siūloma GDR10LT transformavimo schema
Fig. 3.1. Suggested schema of GDR10LT transformation

Pirmasis modelis užtikrins GDR10LT erdviųjų duomenų perkėlimą į atnaujinimo šaltinių duomenų bazę, kurioje bus patikrinti erdviniai duomenys naudojant pradinių duomenų patikros modelius. Šie modeliai turi patikrinti, ar erdviųjų duomenų elementams yra priskirtos reikalingos savybių reikšmės. Jei naudojant patikros modelius bus nustatyti elementų netikslumai, tuomet jie turės

būti patikslinti GDR10LT, jei ne – tuomet duomenys iš atnaujinimo šaltinių duomenų bazės bus generalizuojami naudojant generalizavimo modelius. Generalizuoti erdviniai duomenys bus išsaugomi tarpinių rezultatų duomenų bazėje. Jei transformavimo procese būtų naudojami to paties kartografinio mastelio erdviniai duomenys ir nesiskirtų pradinių ir galutinių erdvių duomenų modeliai, tuomet generalizuoti jų nereikėtų ir užtektų juos perkelti į minėtą duomenų bazę. Sekančiame etape naudojant restruktūrizavimo modelį duomenys saugomi tarpinių rezultatų duomenų bazėje ir pertvarkomi taip, kad atitiktų atnaujinto GDR50LT duomenų modelį. Paskutiniame etape transformuoti erdviniai duomenys yra patikrinami.

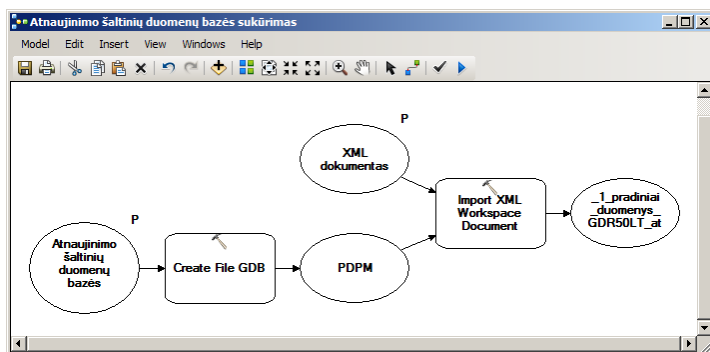
3.2 paveiksle parodyti modeliai sukurti naudojant *ArcGIS* programinės įrangos *ModelBuilder* aplinką. Erdvių duomenų harmonizavimo modeliai sukurti taip, kad galėtų naudoti Lietuvos EII esančias technologijas.

Pirmiausiai sukurtas modelis, skirtas visų GDR10LT erdvių duomenų perkėlimui į suprojektuotą atnaujinimo šaltinių duomenų bazę. Realizavimas pateiktas 3.2 lentelėje.

3.2 lentelė. GDR10LT erdvių duomenų perkėlimo algoritmas

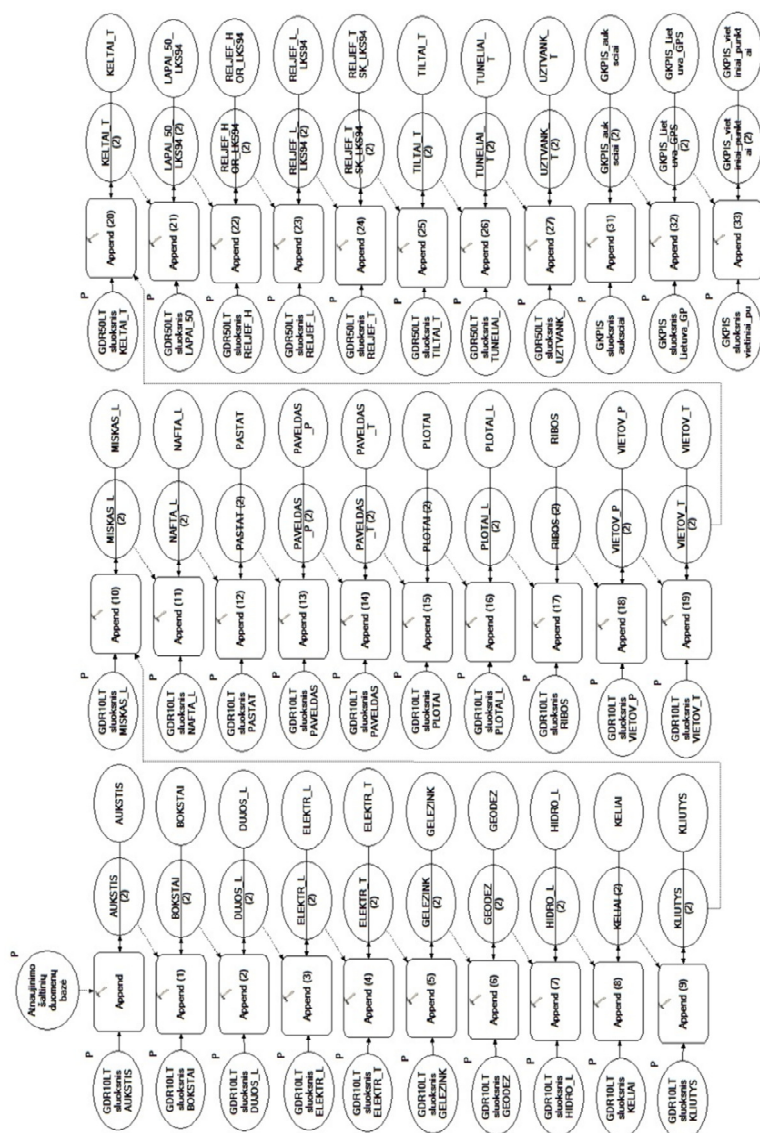
Table 3.2. The algorithm for transfer of spatial data of GDR10LT

Veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Pirmasis modelis: atnaujinimo šaltinių erdvių duomenų bazės sukūrimas naudojant XML saugomą duomenų bazės struktūrą (3.2 pav.).	<i>Create File Geodatabase</i> <i>Import XML Workspace Document</i>
2.	Antrasis modelis: atnaujinimo šaltinių erdvių duomenų įkėlimo į sukurtą atnaujinimo šaltinių duomenų bazę (3.3 pav).	<i>Append</i>



3.2 pav. Atnaujinimo šaltinių duomenų bazės sukūrimo modelio pavyzdys

Fig. 3.2. The model for creating database of data sources



3.3 pav. GDR10LT erdvinių duomenų įkėlimo modelis
Fig. 3.3. Model for uploding of spatial data of GDR10LT

Lentelėje 3.2 aprašytų modelių realizacija naudojant *ArcGIS* programinės įrangos *ModelBuilder* aplinką parodyta 3.2 ir 3.3 paveiksluose.

Paveiksle 3.3 pateiktas modelis perkelia visus GDR10LT erdvinius duomenis į atnaujinimo šaltinių duomenų bazę. Modelis sukurtas taip, kad pradiniai duomenys yra perkeliama ne originalios struktūros, bet į iš anksto aprašytą erdvinių duomenų struktūrą, kuri pakartotiniame GDR10LT ir GDR50LT harmonizavime užtikrintų generalizavimo ir restruktūrizavimo modelių veikimą.

Sekančiame etape sukurti modeliai, skirti GDR10LT erdvinių duomenų patikrai. Atlikus atnaujinimo šaltinių erdvinių duomenų perkėlimą į atnaujinimo šaltinių duomenų bazę, erdviniai duomenys turi būti patikrinti, siekiant rasti erdvinius duomenis ar jų netikslumus, kurie įtakos kitų modelių neveikimą, nekorektišką veikimą ar gautų rezultatų kokybę.

Vykdamas harmonizavimo technologijos kūrimą sukurti trys patikros modeliai, skirti GDR10LT erdvinių duomenų elementų klasių „Keliai“, „Hidro_1“, „Plotai“ erdviniams duomenims patikrinti. Modelių įgyvendinimo pavyzdžiai pateikti C priede.

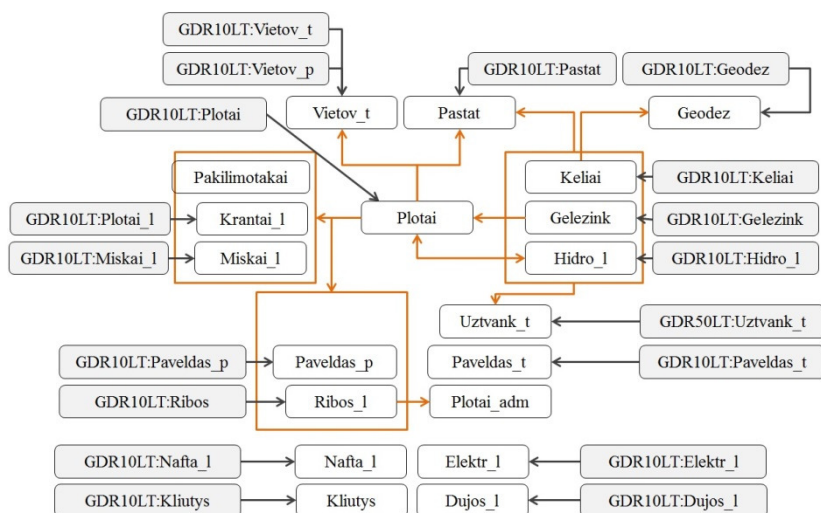
Sukurti modeliai patikrina:

- erdvinių duomenų elementams taikomus topologijos reikalavimus: elementai turi būti vienašaliai, elementai negali perdengti ar kirsti patys savęs, elementai negali perdengti, kirsti ar būti identiški kitiems elementams, esantiems toje pačioje elementų klasėje;
- reikalavimus elementų savybių reikšmėms, kurios turi būti naudojamos atliekant transformavimą.

Atnaujinimo šaltinių erdvinių duomenų struktūra šiuo atveju netikrinama, kadangi numatoma naudoti atnaujinimo šaltinių erdvinius duomenis, kurie jau bus perkelti į atnaujinimo šaltinių duomenų bazę, kurios struktūra užtikrins automatizuotą atnaujinimo modelių veikimą.

Nustačius erdvinių duomenų netikslumus, GDR10LT erdviniai duomenys turėtų būti patikslinti.

Trečiajam transformavimo etapui sukurti generalizavimo modeliai, skirti GDR10LT erdvinių duomenų elementų klasių „Keliai“, „Hidro_1“, „Plotai“ generalizavimui. Kuriant technologiją, skirtą GDR10LT ir GDR50LT harmonizavimui, nustatytas erdvinių duomenų elementų klasių transformavimo eiliškumas, kuris priklauso nuo to, kokie jau generalizuoti GDR50LT erdviniai duomenys turi būti naudojami kitų GDR10LT erdvinių duomenų generalizavime. Paveiksle 3.4 pateikta sąsaja tarp visų harmonizuojamų GDR10LT ir GDR50LT erdvinių duomenų elementų klasių, atsižvelgiant į tai, kurios konkrečios erdvinių duomenų elementų klasės dalyvauja generalizavimo procese.



3.4 pav. Sąsajos tarp GDR10LT ir GDR50LT elementų klasių
Fig. 3.4. The relations between feature classes of GDR10LT and GDR50LT

Paveiksle 3.4 parodyta, kurios GDR10LT erdvinių duomenų elementų klasės turi būti naudojamos generalizavimo procese. Pavyzdžiui, GDR10LT erdvinių duomenų elementų klasės „Vietov_t“ ir „Vietov_p“ turi būti naudojamos transformuojant duomenis į GDR50LT „Vietov_t“; GDR10LT „Plotai“ naudojami transformuojant duomenis į GDR50LT erdvinių duomenų klases „Plotai“, „Pakilimotakai“, o jau transformuoti erdviniai duomenys turi būti papildomai naudojami kartu su GDR10LT „Plotai_l“, juos transformuojant į „Krantai_l“. Tuo tarpu transformavimo metu generalizuojant erdvinių duomenų elementų klasę „Plotai“ papildomai turi būti naudojami jau transformuoti GDR50LT „Keliai“, „Hidro_l“ ir „Gelezink“ ir pan. Remiantis tokiomis aprašytais sąsajomis nustatytas visų GDR10LT elementų klasių generalizavimo eiliškumas, harmonizuojant juos su GDR50LT. Eiliškumas pateiktas 3.3 lentelėje.

3.3 lentelė. Siūlomas elementų klasių generalizavimo eiliškumas
Table 3.3. Proposed priority for generalisation of feature classes

Eil. Nr.	GDR50LT sluoksnio pavadinimas	Generalizavimo proceso vykdymo galimybės
1.	Bokstai	Procesai gali vykti lygiagrečiai vienas su kitu.
2.	Geodez	
3.	Vietov_t	
4.	Elektr_l	

3.3 lentelės pabaiga

Eil. Nr.	Atnaujinamo GDR50LT sluoksnio pavadinimas	Generalizavimo proceso vykdymo galimybės
5.	Dujos_1	
6.	Nafta_1	
7.	Pakilimotakai_1	
8.	Gelezink	
9.	Keliai	
10.	Hidro_1	Procesas turi susidėti iš dviejų iteracijų. Antros iteracijos metu naudojami generalizuoti GDR10LT elementų klasės „Plotai“ elementai.
11.	Plotai	Procesas gali vykti tik generalizavus GDR10LT „Keliai“, „Gelezink“, „Hidro_1“.
12.	Pastat	Procesas gali vykti generalizavus GDR10LT „Keliai“, „Gelezink“, „Hidro_1“, „Plotai“.
13.	Krantai	Procesai gali vykti lygiagrečiai vienas su kitu, tačiau tik generalizavus GDR10LT „Plotai“.
14.	Miskai_1	
15.	Ribos_1	
16.	Ribos_adm	Procesas gali vykti tik generalizavus GDR10LT „Ribos_1“ ir „Vietov_t“.

Toks eiliškumo laikymasis generalizavimo procese, leidžia lygiagrečiai vykdyti kelių elementų klasių generalizavimo procesus, išlaikant nuoseklų erdvių duomenų generalizavimą, atsižvelgiant į tai, kokios jau pilnai ar dalinai generalizuotų elementų klasės bus naudojamos kitame generalizavimo procese. Pavyzdžiui GDR10LT elementų klasės „Bokstai“, „Geodez“, „Vietov_t“ ir pan. generalizavimas gali būti vykdomas lygiagrečiai vienas kitam, kadangi jų erdviniai duomenys nėra susiję vienas su kitu.

3.2. Harmonizuojamų georeferencinių erdvių duomenų generalizavimas

Atsižvelgiant į antrajame skyriuje išnagrinėtus generalizavimo algoritmus, taip pat pateiktus specializuotus generalizavimo sprendimus, juos kombinuojant su kitais *ArcGIS* programinės įrangos algoritmais suprojektuoti ir sukurti modeliai, skirti generalizuoti GDR10LT elementų klases „Keliai“, „Hidro_1“, „Plotai“. Sukurtų modelių apibendrinti aprašymai pateikti 3.2 poskyryje. Palyginimui pateikiami Savino (2011) siūlomi apibendrinti sprendimai.

Elementų vaizduojančių kelius generalizavimas yra sudėtingas ir kompleksinis generalizavimo uždavinys, kuris apima ne tik linijinių elementų

suprastinimą, bet turi būti papildomai taikomi specifiniai algoritmai, skirti tokių kelių tinklo elementų, kaip sankryžos, suprastinimui.

Elementų vaizduojančių kelių generalizavimui sukurtas generalizavimo modelis apimantis 3.4 lentelėje nurodytus pagrindinius žingsnius.

3.4 lentelė. Pagrindiniai kelių generalizavimo žingsniai

Table 3.4. The main steps of generalisation of roads

Siūlomi	Savino (2011)
<ul style="list-style-type: none"> – specifinių kelių elementų suprastinimas; – atšakų eliminavimas; – prioretiškumo reikšmės priskyrimas keliams; – prioretiškumo reikšmės priskyrimas gatvėms; – lygiagrečių elementų glaudinimas; – elementų retinimas; – elementų konfigūracijos suprastinimas. 	<ul style="list-style-type: none"> – greitkelių generalizavimas; – dviejų juostų kelių, poilsio aikštelių, nuvažiavimų identifikavimas; – dviejų juostų kelių glaudinimas; – poilsio aikštelių generalizavimas; – kelių sankryžų generalizavimas; – kelių suprastinimas; – kabančių linijinių elementų eliminavimas.

Savino (2011) siūlomas kelių generalizavimas apima įvairius generalizavimo algoritmus įvairioms kelių tinklo generalizavimo problemoms spręsti, tačiau siūlomame sprendime neatsižvelgiama į kelių tinklo specifiškumą miesto ir užmiesčio teritorijose. Disertacijoje siūlomas kelių tinklo generalizavimo procesas sukurtas atsižvelgiant į GDR10LT erdvinių duomenų ypatybes. Šis algoritmas apima 2.2.4 poskyryje pasiūlytą atšakų eliminavimo algoritmą, yra paprastesnis. Kelių tinklo retinimas vyksta atskirai skirtingose teritorijose taikant skirtingus retinimo parametrus.

Elementų generalizavimo apibendrinta veiksmų seka ir naudojamos pagrindinės *ArcGIS* funkcijos pateikiamos 3.5 lentelėje. Prieš generalizavimą nustatyti kelių tinklo elementų prioritetai, papildomai identifikuojant, kurie elementai negali būti panaikinti. Elementų prioritetai yra naudojami naikinant elementus.

3.5 lentelė. Kelių tinklo generalizavimo seka

Table 3.5. Steps of generalisation of roads network

Apibendrinta veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Atliekama elementų atranka pagal kokybinius parametrus.	<i>Select by Attribute</i>

3.5 lentelės pabaiga

Apibendrinta veiksmų seka:		Pagrindinės ArcGIS funkcijos
2.	Atliekamas specifinių kelių elementų suprastinimas.	<i>Collapse Road Detail</i>
3.	Atliekamas kelių tinklo atšakų tyrimas vertinant bendrą atšakų linijų ilgį. Išskiriamos naikintinos kelių tinklo dalys.	<i>Add Field, Calculate Field, Select Layer By Attribute, Dissolve, Make Feature Layer, Feature Vertices To Points, Select Layer By Location.</i>
4.	Elementams, vaizduojantiems kelius, priskiriama kelių prioriteto reikšmė visame kelių tinkle.	<i>Add Field, Calculate Field.</i>
5.	Elementams, vaizduojantiems gatves, priskiriama gatvių prioriteto reikšmė visame kelių tinkle.	<i>Add Field, Calculate Field.</i>
6.	Suprastinama elemento geometrija (panaikinami kompleksiški elementai: apskritimai, lankai ir pan.)	<i>Densify, Simplify Line, Multipart To Single part, Repair Geometry.</i>
7.	Elementai suskaidomi pagal viršūnių taškus.	<i>Split Line At Vertices</i>
8.	Atliekamas dviejų juostų kelių glaudinimas.	<i>Merge Divided Roads</i>
9.	Atliekamas suskaidytų elementų apjungimas naudojant jiems priskirtas savybių reikšmes.	<i>Dissolve</i>
10.	Atliekamas kelių ir gatvių retinimas visame kelių tinkle atsižvelgiant į mažiausią leidžiamą elemento ilgį ir prioriteto reikšmes. Žingsnis atliekamas keliomis iteracijomis taikant skirtingus parametrus.	<i>Thin Road Network</i>
11.	Pakartotinai atliekamas specifinių kelių elementų suprastinimas.	<i>Collapse Road Detail</i>
12.	Papildomai gali būti atliekamas linijinių elementų suprastinimas taikant Douglas-Peucker algoritmą.	<i>Simplify Line</i>

Kelių tinklo generalizavimo modelio įgyvendinimo pavyzdys pateiktas C priede.

Analizuojant hidrografijos erdvinį duomenų elementus stambesniame mastelyje nustatyta, kad jie gali būti vaizduojami naudojant linijinius arba plotinius elementus. GDR10LT visi hidrografinio tinklo objektai yra išreikšti linijomis, tuo tarpu kitose šalyse kaip pavyzdžiui olandų TOP10NL šie objektai yra vaizduojami tik plotiniais objektais. Tokiu atveju prieš generalizuojant linijinius elementus, būtina taikyti algoritmus, kurie sukurtų linijinius elementus iš plotinių. Uždaviniui išspręsti turi būti taikomas siūlomas modelis, aprašytas 2.2.5 poskyryje.

Analizuojant GDR10LT hidrografijos objektus vaizduojančius elementus nustatyta, kad elementai gali būti suvesti ne upės tekėjimo kryptimi. Todėl sukurtas modelis generalizuoja elementus, neatsižvelgdamas į hidrografijos tinklo logiką. Vis tik reikia pripažinti, kad atlikus tyrimus, nustatyta, kad generalizuojant hidrografijos elementus, sujungtus į nuoseklų logišką tinklą gaunamas patikimesnis rezultatas.

Elementų vaizduojančių hidrografijos tinklą generalizavimui siūloma naudoti generalizavimo modelį, apimantį 3.6 lentelėje nurodytus pagrindinius žingsnius.

3.6 lentelė. Pagrindiniai hidrografijos tinklo generalizavimo žingsniai

Table 3.6. The main steps of generalisation of hydrography network

Siūlomi	Savino (2011)
<ul style="list-style-type: none"> – atranka; – atšakų eliminavimas; – kanalų retinimas; – senvagių eliminavimas; – suprastinimas ir glotninimas. 	<ul style="list-style-type: none"> – matuojamas upių plotis; – sukuriami plotinių elementų centrinės linijos, vaizduojančios upes; – suprastinami linijiniai elementai; – eliminuojami linijiniai elementai, turintys mažesnę ilgį, nei nustatytas minimumas.

Elementų generalizavimo apibendrinta veiksmų seka ir naudojamos *ArcGIS* funkcijos pateikiamos 3.7 lentelėje. Savino (2011) siūloma veiksmų seka pritaikyta generalizuoti GDR10LT elementams ir papildyta veiksmų seka, eliminuojančia senvages.

3.7 lentelė. Hidrografijos tinklo generalizavimo seka

Table 3.7. Steps of generalisation of hydrography network

Apibendrinta veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Atliekama elementų atranka pagal kokybinius parametrus.	<i>Select by Attribute</i>
2.	Nustatomas kiekvieno elemento reikšmingumas, pagal iš anksto nustatytas taisykles, skirtas nustatyti hidrografinio objekto svarbą visame hidrografijos tinkle. Šis parametras apsprendžia elemento prioritetą tinkle.	<i>Add Field, Calculate Field, Make Feature Layer, Select Layer By Location, Feature Class To, Feature Class.</i>
3.	Pirmoji iteracija. Panaikinamos linijinių elementų atšakos, trumpesnės, nei nustatytas leidžiamas mažiausias linijos ilgis.	<i>Add Field, Calculate Field, Select Layer By Attribute, Dissolve, Make Feature Layer, Feature Vertices To Points,</i>

3.7 lentelės tęsinys

Apibendrinta veiksmų seka:		Pagrindinės ArcGIS funkcijos
		<i>Select Layer By Location.</i>
4.	Antroji iteracija. Panaikinamos linijinių elementų atšakos, esančios toliau, nei nustatytas atstumas, nuo elemento turinčio didesnę prioritetą reikšmę.	<i>Add Field, Calculate Field, Select Layer By Attribute, Make Feature Layer, Dissolve, Spatial Join, Feature Vertices To Points, Select Layer By Location, Buffer, Intersect, Split Line At Point, Get Count, Identify, Delete Features.</i>
5.	Trečioji iteracija. Panaikinamos linijinių elementų atšakos, trumpesnės, nei nustatytas leidžiamas mažiausias linijos ilgis, esantis toliau nei nustatytas atstumas, nuo didesnę prioritetą reikšmę turinčių elementų.	<i>Add Field, Calculate Field, Dissolve, Feature Vertices To Points, Spatial Join, Buffer, Make Feature Layer, Select Layer By Attribute, Select Layer By Location, Delete Features.</i>
6.	Ketvirtoji iteracija. Panaikinami linijiniai elementai, trumpesni, nei nustatytas leidžiamas mažiausias linijos ilgis, esantys toliau nei nustatytas atstumas, nuo didesnę prioritetą reikšmę turinčių elementų.	<i>Add Field, Calculate Field, Make Feature Layer, Select Layer By Attribute, Dissolve, Feature To Polygon, Buffer, Select Layer By Location, Intersect, Split Lines At Points, Copy Features, Feature Vertices To Points, Delete Features.</i>
7.	Prieš tai buvęs žingsnis kartojamas penkis kartus, pasirenkant skirtingus parametrus.	–
8.	Panaikinami kabantys linijiniai elementai, esantys nustatyto dydžio uždaruose poligonuose.	<i>Feature To Polygon, Make Feature Layer, Select Layer By Attribute, Dissolve, Feature Vertices To Points, Select Layer By Location, Add Field, Calculate Field.</i>
9.	Vykdomas elementų suprastinimas taikant praplėstą Wang algoritmą.	<i>Simplify Line</i>

Plotinių elementų generalizavimas yra sudėtingas procesas, bet jo apimtis priklauso nuo pradinio ir galutinio erdvinio duomenų modelio ir kaupiamų elementų detalumo ir pan. Pavyzdžiui, Savino (2011) generalizavimo procese siūlo naudoti suprastinimą, agregavimą, kraštinių prailginimą ir suglaudimą.

Tuo tarpu išanalizavus GDR10LT ir GDR50LT, nustatyta, kad generalizuojant erdvinį duomenų elementus reikalinga sukurti modelį įtraukiantį eilę generalizavimo metodų: atranką, agregavimą, suprastinimą, taikant skirtingus algoritmus, pritraukimą prie susijusių elementų, mažų elementų eliminavimą ir pan. Todėl galima teigti, kad žemės dangos generalizavimas tiesiogiai priklauso nuo galutinių erdvinį duomenų rinkinio poreikio.

Paruošto **elementų, vaizduojančių žemės dangą, plotinius hidrografijos objektus ir užstatytas teritorijas, generalizavimo** apibendrinta veiksmų seka ir naudojamos *ArcGIS* funkcijos pateikiamos 3.8 lentelėje. GDR10LT erdvinį duomenų klasės „Plotai“ yra išskirtiniai tuo, kad visų kelių, upių ir geležinkelių dengiamai teritorijai pavaizduoti yra sukurti buferiai aplink atitinkamus elementus GDR10LT elementų klasėse „Keliai“, „Hidro_I“ ir „Gelezink“. Todėl žemės dangos teritorijos yra suskaidytos ir nutolusios viena nuo kitos. Dėl šios priežasties, vienas iš svarbiausių uždavinių generalizuojant „Plotai“ elementus, yra panaikinti buferio dengiamą teritoriją. Šiam uždaviniui spręsti pasirinktas 2.6 lentelėje aprašytas algoritmas.

3.8 lentelė. Žemės dangos generalizavimo seka

Table 3.8. Steps of generalisation of land cover

Apibendrinta veiksmų seka:		Pagrindinės <i>ArcGIS</i> funkcijos
1.	Taikomas 2.6 lentelėje aprašytas algoritmas.	
2.	Atliekama elementų atranka pagal kokybinius ir kiekybinius parametrus, atitinkamai perkeliant atrinktus elementus į atskirus sluoksnius pagal tai kokius gamtinius ar antropogeninius objektus jie vaizduoja.	<i>Make Feature Layer, Select by Attribute, Calculate Field, Feature Class to Feature Class, Dissolve.</i>
3.	Atskirose elementų klasėse agreguojami elementai, vaizduojantys gamtinius objektus, taikant skirtingus agregavimo atstumus.	<i>Aggregate Polygons, Add Field, Calculate Field.</i>
4.	Atskirose elementų klasėse agreguojami elementai, vaizduojantys antropogeninius objektus, taikant skirtingus agregavimo atstumus ir išlaikant ortogonalinę elemento konfigūraciją.	<i>Aggregate Polygons, Add Field, Calculate Field.</i>
5.	Atskirose elementų klasėse suprastinami elementai vaizduojantys gamtinius objektus, taikant skirtingas suprastinimo tolerancijos reikšmes. Taikomas praplėstas Wang algoritmas. Suprastinti objektai apjungiami į vieną sluoksnį.	<i>Simplify Polygons, Repair Geometry, Dissolve, Union, Make Feature Layer, Calculate Field.</i>

3.8 lentelės pabaiga

Apibendrinta veiksmų seka:		Pagrindinės ArcGIS funkcijos
6.	Atskirose elementų klasėse suprastinami elementai vaizduojantys antropogeninius elementus, taikant skirtingas suprastinimo tolerancijos reikšmes. Elementų suprastinimas vykdomas keliomis iteracijomis, vis padidinant tolerancijos reikšmę, tol kol pasiekama maksimali tolerancijos reikšmė.	<i>Simplify Buildings, Repair Geometry, Dissolve, Union, Make Feature Layer, Calculate Field.</i>
7.	Suprastinti elementai apjungiami į vieną sluoksnį.	<i>Append</i>
9.	Keliomis iteracijomis eliminuojami maži elementai, juos prijungiant prie kaimyninių, didinant mažiausią leidžiamą elemento plotą.	<i>Make Feature Layer, Eliminate.</i>

Šio generalizavimo modelio įgyvendinimo pavyzdys pateiktas C priede.

Atliekant modelio eksperimentinius tyrimus, nustatyta, kad elementų vaizduojančių antropogeninius objektus, tokius, kaip užstatytos teritorijos, suprastinimas turi būti vykdomas keliomis iteracijomis keičiant tolerancijos reikšmes (kiekvienoje iteracijoje reikšmė padidinama). Taikant tokį suprastinimo algoritmą yra gaunamas geresnis elemento kartografinis vaizdas. Kiekvienoje iteracijoje padidinant leidžiamą toleranciją, išlieka apie 85 % pradinių duomenų informacijos, o taikant iškart maksimalią leidžiamą toleranciją – apie 75 %.

Sukurtas generalizavimo modelis, lyginant su Savino (2011) siūlomais sprendimais ir nagrinėjamais generalizavimo algoritmais yra išskirtinis galimybe taikyti skirtingus suprastinimo ir agregavimo algoritmus ir tolerancijos reikšmes, priklausomai nuo to, kokią realaus pasaulio objektą vaizduoja elementas, taip pat įdiegtas alternatyvus sprendimas nutolusių elementų agregavimui ir plotinių elementų eliminavimui.

3.3. Harmonizavimo modelio testavimas

Poskyryje 2.1.1 pasiūlyta erdvinių duomenų harmonizavimo metodika pritaikyta erdvinių duomenų rinkinių GDR10LT ir GDR50LT harmonizavimui, transformacijos metu taikant generalizavimo metodus. Sukurtų modelių testavimui buvo naudoti 3.5 paveiksle pavaizduota GDR10LT erdviniai duomenys dengiantys apie 500 tūkst. ha teritoriją.

Modelių testavimai buvo atliekami dviem etapais:

1. sudarant GDR10LT elementų klasių „Keliai“, „Hidro_l“ ir „Plotai“ generalizavimo algoritmus ir lygiagrečiai realizuojant juos harmonizavimo modeliuose, siekiant parinkti tinkamiausius *ArcGIS* įrankius, procesų eiliškumą, parametrus ir procesų iteracijų skaičių, kaip siūloma erdvinių duomenų harmonizavimo metodikoje;
2. atliekant sudarytų modelių testavimą, norint įsitikinti jų teisingumu. Šiame etape buvo atliekamas GDR10LT testuojamoje teritorijoje erdvinių duomenų transformavimas, taikant erdvinių duomenų generalizavimo algoritmus.

Kelių tinklo, hidrografijos tinklo ir žemės dangos erdvinių duomenų pokyčiai, atlikus generalizavimą pateikti 3.9–3.11 lentelėse.

3.9 lentelė. Kelių tinklo elementų pokyčių analizės rezultatai

Table 3.9. The results of analysis of changes in features of roads network

GKODAS	Buvęs kiekis	Kiekis po gen.	Pokytis	Pokytis %	Buvęs ilgis, m	Ilgis po gen., m	Pokytis, m	Pokytis %
gc12	2629	2002	627	23,85	735682	707788	27894	3,79
gc12gc2	2850	2177	673	23,61	582632	575275	7358	1,26
gc14	5917	4780	1137	19,22	2667211	2634878	32333	1,21
gc14gc2	553	427	126	22,78	155472	155422	50	0,03
gc15	627	510	117	18,66	341980	334892	7087	2,07
gc16	9295	6024	3271	35,19	3257688	3003954	253734	7,79
gc2	8616	5324	3292	38,21	1520006	1267481	252525	16,61

3.10 lentelė. Kelių tinklo elementų pokyčių analizės rezultatai

Table 3.10. The results of analysis of changes in features of roads network

GKODAS	Buvęs kiekis	Kiekis po gen.	Pokytis	Pokytis %	Buvęs ilgis, m	Ilgis po gen., m	Pokytis, m	Pokytis %
fhc3	253	241	12	4,74	42892	41256	1636	3,82
hc1	87	87	0	0	80100	77300	2800	3,50
hc3	8	8	0	0	2459	2455	4	0,16
hc31	515	442	73	14,17	161992	143915	18076	11,16

3.10 lentelės pabaiga

GKODAS	Buvęs kiekis	Kiekis po gen.	Pokytis	Pokytis %	Buvęs ilgis, m	Ilgis po gen., m	Pokytis, m	Pokytis %
hc31op0	136	128	8	5,88	2730	2587	143	5,24
hc32	1084	949	135	12,45	358761	324367	34394	9,59
hc32op0	243	230	13	5,35	3803	3626	176	4,64
hc33	889	857	32	3,6	330325	313710	16615	5,03
hc33op0	340	334	6	1,76	5541,59	5450,66	90	1,64
hc3op0	1	1	0	0	18,43	18,43	0	0

3.11 lentelė. Žemės dangos elementų pokyčių analizės rezultatai**Table 3.11.** The results of analysis of changes in land covers features

GKO-DAS	Buvęs kiekis	Kiekis po gen.	Pokytis	Pokytis, %	Buvęs plotas, m ²	Plotas po gen., m ²	Pokytis, m ²	Pokytis, %
hd1	42	34	8	19,05	12613873	12767491	-153618	-1,22
hd2	86	63	23	26,745	3323387	3308884	14503	0,44
hd3	60	26	34	56,67	4427547	4399198	28349	0,64
hd4	7439	320	7119	95,7	8441478	5203561	3237915	38,36
hd5	2	2	0	0	604704851	604733181	-28330	-0,004
hd6	304	103	201	66,12	12342769	11859903	482866	3,91
hd9	86	73	13	15,12	9100538	9132540	-32002	-0,35
ms0	7028	2831	4197	59,72	1305161757	1397994001	-92832252	-7,11
ms4	298	142	156	52,35	16533090	13664164	2868925	17,35
pu0	14457	5584	8873	61,38	154393119	168616997	-14223878	-9,21
pu3	1505	723	782	51,96	32021107	30340585	1680522	5,24
sd42	55	37	18	32,73	12867052	12887811	-20759	-0,16
va1	7	5	2	28,57	1920795	2005715	-84920	-4,42
vp1	364	99	265	72,8	1964787	1851989	112798	5,74
gt12	754	0	754	N/D	59266091	0	59266091	N/D
gt14	1507	0	1507	N/D	25403355	0	25403355	N/D
gt15	314	0	314	N/D	2228313	0	2228313	N/D
gt16	5910	0	5910	N/D	17571081	0	17571081	N/D

Lentelės 3.11 pabaiga

GKO-DAS	Buvęs kiekis	Kiekis po gen.	Pokytis	Pokytis, %	Buvęs plotas, m ²	Plotas po gen., m ²	Pokytis, m ²	Pokytis, %
gt18	97	0	97	N/D	1369984	0	1369984	N/D
gt19	96	0	96	N/D	528320	0	528320	N/D
gt2	1060	0	1060	N/D	12604182	0	12604182	N/D
hd21	2598	0	2598	N/D	2999566	0	2999566	N/D
hd22	4888	0	4888	N/D	13574434	0	13574434	N/D
hd23	4265	0	4265	N/D	46711549	0	46711549	N/D
sd11	10016	0	10016	N/D	3249649558	0	3.25E+09	N/D
sd15	10273	0	10273	N/D	66354919	0	66354919	N/D
sd2	11663	0	11663	N/D	121105383	0	121105383	N/D
sd4	3145	0	3145	N/D	28850976	0	28850976	N/D
sd7	204	0	204	N/D	16197656	0	16197656	N/D
va11	3	0	3	N/D	52452	0	52452	N/D
vk1	174	0	174	N/D	1524790	0	1524790	N/D

Lentelėje naudojama reikšmė N/D – reiškia, kad reikšmė nenustatyta, kadangi atlikus generalizavimą, elementai buvo eliminuoti.

Atlikus generalizuotų elementų, vaizduojančių kelius, analizę, nustatyta, kad sumažėjo elementų kiekis ir jų ilgis. Elementai, kurių GKODAS reikšmė „gc2“ pasikeitė daugiausiai: kiekis apie 38 %, o ilgis virš 16 %. Šie elementai vaizduoja gatves, kurių tankumas gyvenamosiose teritorijose yra didelis, nei kelių užmiesčio teritorijose. Dėl to tarp jų yra mažesni atstumai. Todėl harmonizuojant erdvinius duomenis tarp skirtingų mastelių, atsižvelgiant į galimus mažiausius atstumus tarp objektų šie elementai yra išretingami labiau, nei kiti. Generalizuojant kelių tinklą yra eliminuojamos per trumpos elementų atšakos, kurios nėra vizualiai matomos smulkesnio mastelio žemėlapiuose. Atliekant GDR10LT elementų vaizduojančių kelius eliminavimą, panaikintos atšakos, kurių ilgis nesiekė 200 m užmiestyje ir 100 m miestų teritorijose. Todėl 7,79 % sumažėjo elementų vaizduojančių lauko ir miško keliukus (GKODAS reikšmė „gc16“), kadangi GDR10LT tokie elementai dažnai vaizduoja įvažiavimus į sodybas, kurie GDR50LT yra per trumpi ir dėl to nevaizduojami.

Generalizavus erdvinių duomenų elementų klasę „Hidro_I“, kaip ir generalizuojant elementų klasę „Keliai“, sumažėjo elementų kiekis ir ilgis. Taikant siūlomą hidrografijos tinklo elementų generalizavimo algoritmą

daugiausiai buvo panaikinta elementų turinčių GKODAS reikšmę „hc31“. Jų sumažėjo kiekis sumažėjo apie 14 %, o ilgis apie 11 %.

Elementų, vaizduojančių žemės dangą, pokyčiai lyginant su kelių tinklo elementų pokyčiais, atlikus generalizavimą yra kitokie. Jei elementų, vaizduojančių kelius, kiekis ir ilgis atlikus generalizavimą visuomet sumažėja, tai generalizuojant plotinius elementus kiekis sumažėja, tačiau dengiamas plotas gali padidėti. Elementų plotas gali padidėti dėl:

1. eliminuojamų elementų, kurių plotu padidėja kaimyniniai plotiniai elementai, tai yra svarbesni elementai agreguoja ne tokius svarbius elementus;
2. taikomų elementų ribos suprastinimo algoritmų, kurie eliminuoja nereikšmingus vingius.

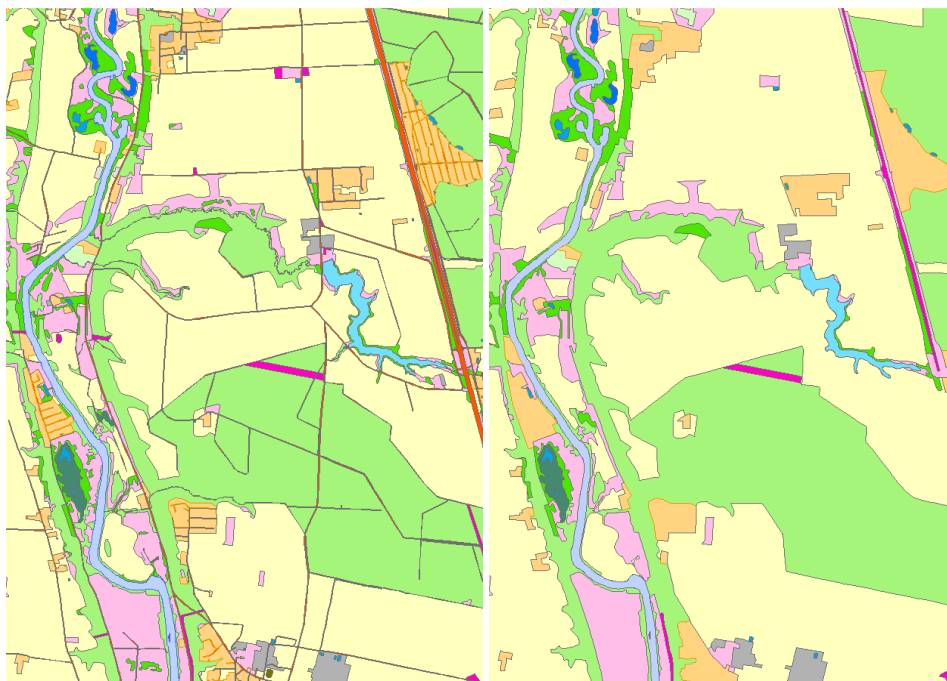
Analizuojant elementų, vaizduojančių žemės dangą, pokyčius prieš ir po generalizavimo nustatyta, kad labiausiai sumažėjo elementų, turinčių GKODAS reikšmę „hd4“ kiekis ir bendras užimamas plotas. Tuo tarpu nors elementų turinčių GKODAS reikšmę „ms0“ kiekis po generalizavimo sumažėjo, tačiau jų plotas padidėjo 7,11 %, kadangi buvo agreguoti nereikšmingi kaimyniniai elementai.

Paprastai generalizuotų erdvinių duomenų kartografinis vaizdas turi atitikti tam masteliui taikomus tikslumo reikalavimus. Juos galima pasiskaičiuoti taikant 2.20 formulę. Tačiau vertinant erdvinių duomenų rinkinio elementus duomenys turi būti detalesni, tam, kad tenkintų skirtingų vartotojų poreikius.

Vertinant erdvinių duomenų transformavimo teisingumą yra svarbiau išanalizuoti gautų erdvinių duomenų elementų vaizdo korektiškumą ir atitikimą galutiniam erdvinių duomenų modeliui. Gautų erdvinių duomenų bendro vaizdo pavyzdys pateiktas 3.5–3.9 paveiksluose. Analizuojant vaizdą, nustatyta, kad generalizuoti erdviniai duomenys tenkina pagrindinius reikalavimus, tačiau siekiant gautų duomenų tikslumo, reikėtų didinti testuojamą teritoriją, ir prireikus generalizavimo algoritmus tikslinti, priklausomai nuo erdvinių duomenų specifiškumo skirtingose teritorijose.

Gautų rezultatų teisingumui įvertinti sukurti patikros modeliai, kurie patikrina gautų po generalizavimo erdvinių duomenų klasių „Plotai“, „Keliai“, „Hidro_l“ elementus, ar:

- gauti rezultatai atitinka taikomus jiems topologinius reikalavimus: ar elementai yra vienadaliai ir nedengia vienas kito, ar tarp plotinių elementų nėra tarpų;
- priskirtos elementų savybių reikšmės yra teisingos.



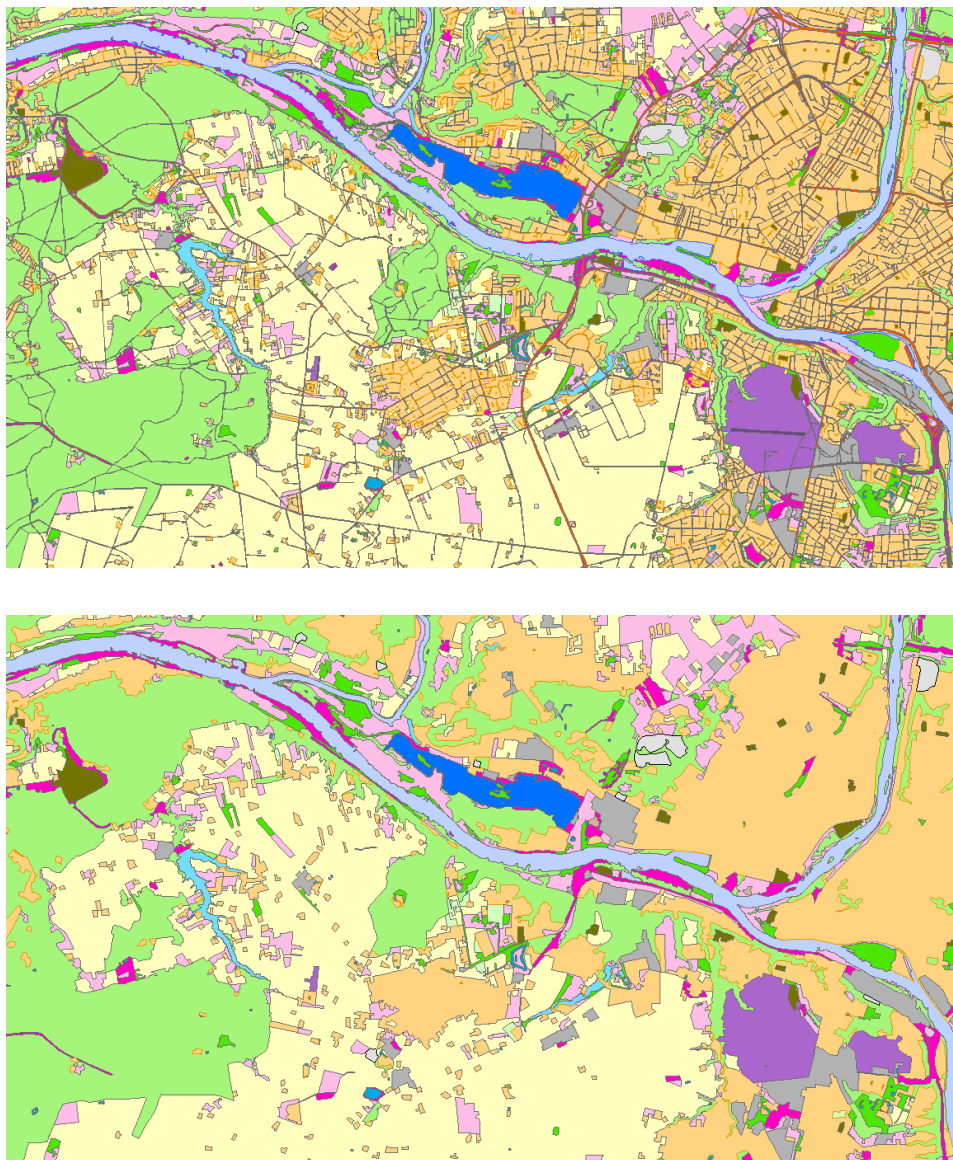
3.5 pav. Pavyzdys GDR10LT erdvinių duomenų elementų klasės „Plotai“ prieš (kairėje) ir po (dešinėje) generalizavimo (elementų agregavimas)

Fig. 3.5. The example of GDR10LT features class “Plotai” before (left) and after (right) generalisation (aggregation of features)



3.6 pav. Pavyzdys GDR10LT erdvinių duomenų elementų klasės „Plotai“ prieš (kairėje) ir po (dešinėje) generalizavimo pavyzdys (elementų agregavimas ir konfigūracijos suprastinimas)

Fig. 3.6. The example of GDR10LT features class “Plotai” before (left) and after (right) generalisation (aggregation of features and simplification of shape)

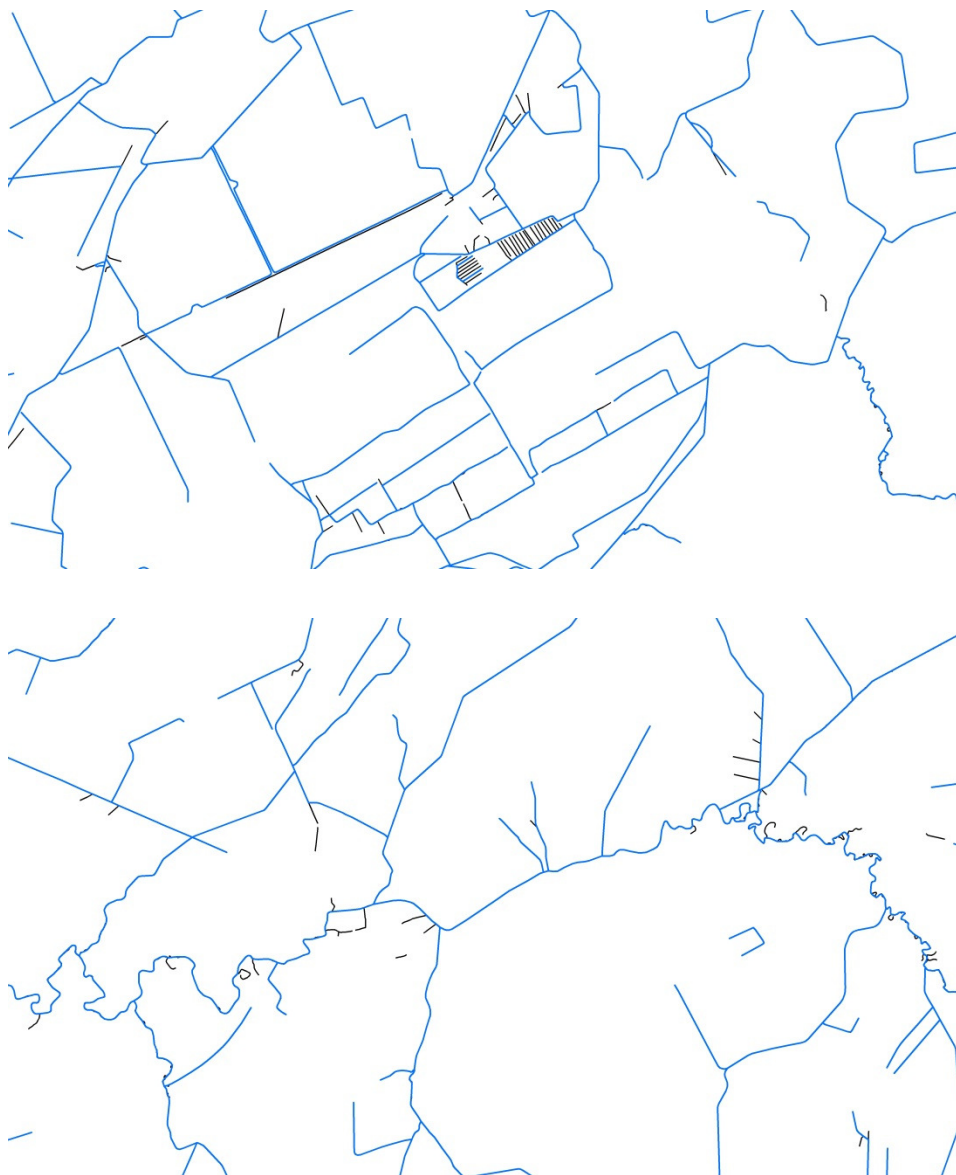


3.7 pav. Pavyzdys GDR10LT erdvinių duomenų elementų klasės „Plotai“ prieš (viršuje) ir po (apačioje) generalizavimo

Fig. 3.7. The example of GDR10LT features class “Plotai” before (above) and after (below) generalisation



3.8 pav. Pavyzdys GDR10LT erdvinių duomenų elementų klasės „Keliai“ prieš (pavaizduota raudonai) ir po (pavaizduota juodai) generalizavimo
Fig. 3.8. The example of GDR10LT features class “Keliai” before (represented in red) and after (represented in black) generalisation



3.9 pav. Pavyzdys GDR10LT erdvinių duomenų elementų klasės „Hidro_1“ prieš (pavaizduota juodai) ir po (pavaizduota mėlynai) generalizavimo
Fig. 3.9. The example of GDR10LT features class “Hidro_1” before (represent in black) and after (represent in blue) generalisation

Patikros modeliai buvo taikomi galutiniams rezultatams įvertinti. GDR50LT erdvinių duomenų patikros rezultatai pateikti 3.11 lentelėje.

3.11 lentelė. Generalizuotų elementų patikros rezultatai
Table 3.11. The results of verification of generalised features

Patikros taisyklė	Generalizuotų elementų klasė:		
	„Keliai“	„Hidro_1“	„Plotai“
Elementai negali persidengti tarpusavyje toje pačioje elementų klasėje.	0	0	3
Tarp elementų negali būti tarpų toje pačioje elementų klasėje.	–	–	15
Elementai negali kirsti patys savęs.	0	0	0
Negali būti pseudo taškų.	0	0	–
Elementai negali būti daugiadaliai.	0	0	0
Elementai negali būti mažesni, nei nustatytas minimumas.	36	45	0
Elementams priskirtos teisingos savybių reikšmės.	0	0	0

Patikros rezultatai parodė, kad atlikus plotinių elementų generalizavimą gali likti persidengimų arba tarpų tarp elementų. Išanalizavus duomenis, nustatyta, kad šie netikslumai yra maži. Tai galėjo atsitikti dėl sudėtingų modelių, kuriuose naudojamas didelis kiekis skirtingų funkcijų arba erdviniams duomenims taikomų tikslumų. Išanalizavus nustatytus netikslumus elementų klasėse „Keliai“ ir „Hidro_1“ nustatyta, kad per trumpi elementai gali atsirasti tais atvejais, kaip elementų pradžios ir pabaigos taškuose susieina daugiau nei du linijiniai elementai. Tačiau bendrai erdvinių duomenų kokybei tokių trumpų elementų išlaikymas neįtakoja. Prireikus, tokie elementai gali būti eliminuojami sutapdinant jų pradžios ir pabaigos taškus.

3.4. Trečiojo skyriaus išvados

1. Taikant siūlomą erdvinių duomenų harmonizavimo metodiką sukurta ir praktiškai patikrinta technologija, skirta harmonizuoti kelių ir hidrografijos tinklo, bei žemės dangos GDR10LT ir GDR50LT erdvinis duomenis. Technologija sukurta naudojant programinę įrangą identišką veikiančiai Lietuvos erdvinės informacijos infrastruktūroje.
2. Sukurti modeliai, skirti georeferenciniams erdviniams duomenims, vaizduojantiems kelius, upes, žemės dangą generalizuoti. Modeliai gali

būti taikomi transformuojant duomenis tarp harmonizuojamų skirtingo mastelio erdvinio duomenų rinkinių.

3. Siekiant tikslesnio kartografinio vaizdo, sukurtas linijinių elementų, vaizduojančių kelius ir gatves, generalizavimo modelis. Jame priklausomai nuo to, ar elementas vaizduoja kelio, ar gatvės ašinę liniją, taikomi skirtingi kelių retinimo parametrai, kadangi gyvenamųjų vietovių teritorijose gatvių tinklas yra tankesnis, nei kelių, esančių neurbanizuotose teritorijose.
4. Atliekant eksperimentinius tyrimus nustatyta, kad elementų vaizduojančių antropogeninius objektus suprastinimas turi būti vykdomas keliomis iteracijomis keičiant tolerancijos reikšmes (kiekvienoje iteracijoje reikšmę padidinant). Taikant tokį suprastinimo algoritmą yra gaunamas tikslesnis elemento kartografinis vaizdas. Kiekvienoje iteracijoje padidinant leidžiamą toleranciją, išlieka apie 85 % pradinio duomenų informacijos, o taikant iškart maksimalią leidžiamą toleranciją – apie 75 %.
5. Atliekant linijinių elementų generalizavimą sumažėja jų kiekis ir bendras ilgis. Naudojant siūlomus generalizavimo modelius elementams vaizduojantiems kelius ir gatves, galima sumažinti elementų ilgį vidutiniškai iki 5 %, o kiekį apie 25 %. Naudojant šį modelį daugiausiai yra panaikinami trumpi pertekliniai linijiniai elementai.
6. Generalizuojant plotinius elementus, jų plotas gali padidėti dėl nereikšmingų kaimyninių elementų agregavimo arba dėl generalizavime taikomų elementų konfigūracijos suprastinimo algoritmų. Pavyzdžiui, atliekant elementų, vaizduojančių žemės dangą ir vandens telkinius, buvo panaikinami kartografinė prasme maži elementai. Elementų vaizduojančių miškus plotas padidėjo apie 7 %, nors jų kiekis ir sumažėjo.
7. Siūloma harmonizavimo metodika leidžia sukurti tik griežtomis taisyklėmis iš anksto apibrėžtus generalizavimo modelius, kurie kuriami atsižvelgiant į pradinio ir galutinio erdvinio duomenų specifiškumą.

Bendrosios išvados

Remiantis disertaciniame darbe atliktais tyrimais, gautos šios apibendrintos išvados:

1. Taikant siūlomą erdvinių duomenų harmonizavimo metodiką sukurta technologija, skirta harmonizuoti skirtingo mastelio georeferencinius erdvinius duomenis, vaizduojančius žemės dangą, kelių ir hidrografijos tinklą. Eksperimentų metu erdvinių duomenų rinkinių GDR10LT ir GDR50LT harmonizavimo modeliai sukurti naudojant identišką programinę įrangą, kuri veikia Lietuvos erdvinės informacijos infrastruktūroje.
2. Parengta patobulinta erdvinių duomenų harmonizavimo metodika, kurioje numatyta erdvinių duomenų pokyčių nustatymo ir vertinimo procedūra, detalizuotas plotinių elementų konfiguracijos pokyčio įvertinimas.
3. Pasiūlytas erdvinių duomenų modeliavimo patobulinimas, skirtas kaupti informaciją apie erdvinių duomenų elementų pokyčius ir jų pobūdį.
4. Parengti metodai, kuriais nustatomi ir vertinami erdvinių duomenų pokyčiai. Taikant šiuos metodus galima nustatyti potencialias erdvinių duomenų transformavimo vietas, priklausomai nuo galutinio erdvinių duomenų modelio ir erdvinių duomenų kartografinio mastelio.

5. Pasiūlyti erdvinių duomenų pokyčių nustatymo metodai papildomai įvertina erdvinių duomenų plotinių elementų geometrijos pokyčius. Todėl erdvinių duomenų harmonizavimo procese perkelti erdvinius duomenis iš pradinių duomenų šaltinių bazės, prireikus galima perkelti tik tuos elementus, kurių savybių ar geometrijos pokyčiai yra esminiai galutiniam erdvinių duomenų rinkiniui, priklausomai nuo jam taikomo duomenų modelio ir kartografinio mastelio. Tokie metodai įgyvendina dalinį erdvinių duomenų elementų transformavimą ir prireikus generalizavimą erdvinės informacijos harmonizavimo procese.
6. Patobulinti algoritmai erdvinių duomenų elementams eliminuoti ir apjungti, panaikinti linijinių elementų atšakas, sutapdinti kaimyninius elementus, gali būti taikomi pavienių erdvinių duomenų ar erdvinių duomenų rinkiniams generalizuoti, taip pat harmonizuoti skirtingų kartografinių mastelių erdvinių duomenų rinkinius.
7. Kuriant skirtingo mastelio erdvinių duomenų rinkinių harmonizavimo technologijas būtina sudaryti generalizuojamų erdvinių duomenų elementų prioritetinę seką, atsižvelgiant į sąsajas tarp jų.

Literatūra ir šaltiniai

Afflerbach, S.; Illert, A.; Sarjakoski, T. 2004. The Harmonisation Challenge of Core National Topographic Databases in the EU-Project GiMoDig, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 35(B4:4): 129–134.

Ai, T. H. 2004. A generalization of contour line based on the extraction and analysis of drainage system, *International Archive of Photogrammetry and Remote Sensing* 35(4/3).

Armenakis, C.; Cyr, I.; Papanikolaou, E. 2002. Change detection methods for the revision of topographic databases, *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences* 34(4): 792–797.

Badard, T. 1998. Towards a generic updating tool for geographic databases, *GIS/LIS'98* 352–363.

Badard, T. 1999. On the automatic retrieval of updates in geographic databases based on geographic data matching tools, *Bulletin du Comité français de cartographie* 162: 34–40.

Badard, T.; Lemarie, C. 1999. Propagating updates between geographic databases with different scales, *Innovation in GIS VII: GeoComputation* 135–146.

Bader, M.; Barrault, M. 2000. Improving Snakes for linear feature displacement in cartographic generalization, *GeoComputation*. [interaktyvus] [žiūrėta 2012 m. kovo 23 d.]. Prieiga per internetą: <http://www.geocomputation.org/2000/GC034/Gc034.htm>

Baella, B.; Pla, M. 1999. Map generalization to obtain the topographic map of Catalonia, *ICA Workshop on Progress in Automated Map Generalisation*. [interaktyvus] [žiūrėta 2013 m. gruodžio 21 d.]. Prieiga per internetą:

http://www.icc.es/index.php/esl/content/download/3603/11698/file/map_generalization_obtain_mtc10m.pdf

Baella, B.; Pla, M. 2003. An example of database generalization workflow: the Topographic Database of Catalonia at 1:25.000, *ICA Workshop on Map Generalization*. [interaktyvus] [žiūrėta 2013 m. gruodžio 21 d.]. Prieiga per internetą: http://www.icc.cat/index.php/content/download/3586/11650/file/example_database_generalization_workflow.pdf.

Baella, B.; Pla, M. 2005. Reorganizing the Topographic Databases of the Institut Cartogràfic de Catalunya applying generalization, *8th ICA Workshop on Generalization and Multiple Representation*. [interaktyvus] [žiūrėta 2013 m. gruodžio 21 d.]. Prieiga per internetą:

http://www.icc.es/index.php/cat/content/download/3582/11634/file/reorganizing_topographic_databases_icc.pdf.

Balboa, J. L. G.; Lopez, F. J. A. 2000. Frequency filtering of linear elements by means of wavelets – A method and example, *The Cartographic Journal* 37(1): 39–50.

Beard, K. 1987. How to survive on a single detailed database. *Proceedings Auto–Carto 8*: 211–220.

Beard, M. K.; Mackaness, W. 1991. Generalization operators and supporting structures. *Proceedings Auto–Carto 10*: 29–45.

Beconytė, G.; Kryžanauskas, A.; Papšienė, L.; Papšys, K.; Stankevičius, Ž. 2009. Lietuvos geografinės informacijos infrastruktūra – kelias į bendrą geografijos metodologiją, *Geografija* 45(1): 1–10.

Bergenheim, W.; Sarjakoski, L. T.; Sarjakoski, T. 2009. A Web processing service for GRASS GIS to provide on-line generalisation, *Proceedings of the 12th AGILE International Conference on Geographic Information Science (AGILE)*.

De Berg, M.; van Kreveld, M.; Schirra, S. 1998. Topologically correct subdivision simplification using the bandwidth criterion, *Cartography and Geographic Information Systems* 25(4): 243–257.

Bobzien, M.; Burghardt D.; Petzold I. 2005. Regeneralization and construction–Two alternative approaches to automated incremental updating in MRDB. *Proceedings of 22nd International Cartographic Conference*.

Boutoura, C. 1989. Line generalization using spectral techniques. *Cartographica* 26(3&4): 33–48.

Brassel, K.; Weibel R. 1988. A review and conceptual framework of automated map generalization. *International Journal of Geographical Information Science* 2(3): 229–244.

Braun, A.; Halbecq, X.; Lecordix, F.; Le Gallic, J. M.; Prigent, F. 2007. A new flowline for the French topographic maps in IGN, *Proceedings of the 22nd International Cartographic Conference*.

Brewer, C.A.; Battenfield, B.P. 2007. Framing Guidelines for Multi-Scale Map Design Using Databases at Multiple Resolutions, *Cartography and GIS* 34(1): 3–15.

Burghardt, D. 2005. Controlled Line Smoothing by Snakes, *GeoInformatica* 9: 237–252.

Burghardt, D.; Meier, S. 1997. Cartographic displacement using the snakes concept, *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps* 59–71.

Butenuth, M.; Gosseln, G.; Tiedge, M.; Heipke, C.; Lipeck, U.; Sester, M. 2007. Integration of heterogeneous geospatial data in federate database, *ISPRS Journal of Photogrammetry and Remote Sensing* 62(5): 328–346.
<http://dx.doi.org/10.1016/j.isprsjprs.2007.04.003>

Battenfield, B. P. 2002. Transmitting Vector Geospatial Data across the Internet. *Geographic Information Science, Computer Science* 2478: 51–64.

Battenfield, B. P.; McMaster, R. B. 1991. Map generalization: making rules for knowledge representation, *Longman Proceedings of the 8th International Symposium on Spatial Data Handling* 214–224.

Cecconi, A. 2003. *Integration of Cartographic Generalization and Multi-Scale Databases for Enhanced Web Mapping*, PhD thesis, Department of Geography, University of Zurich, 155 p.

Cecconi, A.; Galanda, M. 2002. Adaptive Zooming in Web Cartography, *Computer Graphics Forum* 21(4): 787–799.

Christ, F. 1978. A Program for the Fully Automated Displacement of Point and Line Features in Cartographic Generalizations, *Informations Relative to Cartography and Geodesy, Translations* 35: 5–30.

Collins, M.; Rhind, J. 1997. Developing global environmental databases: lessons learned about framework information, *Framework for the World* 120–129.

Comber, A.; Fisher, P.; Wadsworth, R. 2004. Integrating land-cover data with different ontologies: identifying change from inconsistency, *International Journal of Geographical Information Science* 18(7): 691–708.
<http://dx.doi.org/10.1080/13658810410001705316>

Cromley, R.; Campbell, G. 1990. A Geometrically Efficient Bandwidth Line Simplification Algorithm, *Proceedings of the 4th International Symposium on Spatial Data Handling* 1: 77–84.

De Berg, M.; van Kreveld, M.; Schirra, S. 1998. Topologically correct subdivision simplification using the bandwidth criterion, *Cartography and Geographic Information Systems* 25(4): 243–257.

de Vries, M.; di Donato, P.; Penninga, F. 2007b. A7.1-D1 Concept of application-specific harmonised data models, *Darmstadt: HUMBOLDT Consortium*. [interaktyvus] [žiūrėta 2013 m. spalio 1 d.]. Prieiga per internetą: http://www.esdi-humboldt.eu/files/818-a7_1d1_concept_of_application-specific_harmonised-tud-001-final.pdf

de Vries, M.; Giger, Ch.; Loidold, M. 2007a. A3.5-D1 State of the Art in Data Harmonisation and Data Management. *Darmstadt: HUMBOLDT Consortium*.

Delattre, N. 2004. EuroRegionalMap, a Pragmatic Solution to Create an ESDI at Medium Scale Through the Harmonization of National Topographic Data Collections, *10th EC GI & GIS Workshop, ESDI State of the Art* 23–25.

DeLucia, A. A.; Black, R. B. 1987. A comprehensive approach to automatic feature generalization, *Proceedings of 13th International Cartographic Conference* 4: 169–192.

Desclée, B.; Bogaert, P.; Defourny, P. 2006. Forest change detection by statistical object-based method, *Remote Sensing of Environment* 102(1): 1–11. <http://dx.doi.org/10.1016/j.rse.2006.01.013>

Devegele, T. 1998. *Le processus d'intégration et d'appariement des BD géographiques*, PhD thesis, University Versailles-Saint Quentin, France, 209 p.

Donaubauer, A.; Fichtinger, A.; Schilcher, M.; Straub, F. 2006. Model driven approach for accessing distributed spatial data using web services-demonstrated for cross-border GIS applications, *XXIII International FIG Congress* 8–13. [interaktyvus] [žiūrėta 2011 m. rugsėjo 7 d.]. Prieiga per internetą: https://www.fig.net/pub/fig2006/papers/ts82/ts82_04_donaubauer_et_al_0603.pdf

Douglas D.; Peucker T. 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, *Cartographica, The International Journal for Geographic Information and Geovisualization* 10: 112–122.

Draft Technical Guidance for INSPIRE Coordinate Transformation Service. Version 2.1. 2010a. [interaktyvus] [žiūrėta 2011 m. rugsėjo 7 d.]. Prieiga per internetą: http://inspire.ec.europa.eu/documents/Network_Services/INSPIRE_Draft_Technical_Guidance_Coordinate_Transformation_Services_%28version_2%201%29.pdf

Estes, J.; Lawless, J.; Mooneyhan, D. W. 1995. Report on the International Symposium on Core Data Needs for Environmental Assessment and Sustainable Development Strategies. *New York, United Nations Development Programme/United Nations Environment Programme*

Europos Parlamento ir Tarybos direktyva 2007/2/EB 2007 m. kovo 14 d. sukurianti Europos bendrijos erdvinės informacijos infrastruktūrą (INSPIRE). [interaktyvus] [žiūrėta 2011 m. rugsėjo 7 d.]. Prieiga per internetą: <http://eur-lex.europa.eu/legal-content/LT/TXT/HTML/?uri=CELEX:32007L0002&from=LT>

Feranec, J.; Hazeu, G.; Christensen, S.; Jaffrain, G. 2007. Corine land cover change detection in Europe (case studies of the Netherlands and Slovakia), *Land Use Policy* 24(1): 234–247. <http://dx.doi.org/10.1016/j.landusepol.2006.02.002>

Feringa, W. 2001. File Formats and Plugins, *Web Cartography: Developments and Prospects* 177–193.

Fichtinger, A.; Rix, J.; Schaffer, U.; Michi, I.; Gone, M.; Reitz, T. 2011. Data Harmonisation Put into Practice by the HUMBOLDT Project, *International Journal of Spatial Data Infrastructure Research* 6: 234–260. <http://dx.doi.org/10.2902/1725-0463.2011.06.art11>

Frank, R.; Ester, M. 2006. A quantitative similarity measure for maps, *Progress in Spatial Data Handling* 435–450.

GDR10LT. 2014. *Lietuvos Respublikos teritorijos M1:10 000 georeferencinių erdvinių duomenų rinkinio GDR10LT specifikacija* [interaktyvus] [žiūrėta 2014 m. gegužės 2 d.]. Prieiga per internetą: http://www.geoportal.lt/download/Specifikacijos/GDR10LT_specifikacija.pdf

GDR50LT. 2012. *Lietuvos Respublikos teritorijos M1:10 000 georeferencinių erdvinių duomenų rinkinio GDR10LT specifikacija* [interaktyvus] [žiūrėta 2014 m. gegužės 2 d.]. Prieiga per internetą: http://www.geoportal.lt/download/Specifikacijos/GDR50LT_specifikacija.pdf

Gedrange, C.; Neubert, M.; Rohnert, S. 2011. Cross-border harmonisation of spatial base data between Germany and the Czech Republic, *1 International Journal of Spatial Data Infrastructures Research* 6: 53–72. <http://dx.doi.org/10.2902/1725-0463.2011.06.art3>

Geodezijos ir kartografijos įstatymas 2010. [interaktyvus] [žiūrėta 2013 m. gegužės 5 d.]. Prieiga per internetą: http://eur-lex.europa.eu/legal-content/EN/ALL/;ELX_SESSIONID=CYm1J6xQK2YQ0QygnC5RWG1kTLLGzm1nJHD7qyKN1Szkd5ZDJ1Q!-558024651?uri=CELEX:32007L0002

Giger, C.; Klien, E.; Reitz, T.; Rix, J. 2014. Methodology of Data Harmonisation and Integration in Spatial Data Infrastructures, 290 p.

Gokgoz, T. 2005. Generalization of contours using deviation angles and error bands, *The Cartographic Journal* 42(2): 145–156.

Gosseln, G. V.; Sester, M. 2004: Intergration of geoscientific data sets and the German digital map using a matching approach, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 1249–1254. [interaktyvus] [žiūrėta 2012 m. liepos 21 d.]. Prieiga per internetą: <http://www.isprs.org/proceedings/XXXV/congress/comm4/papers/534.pdf>

Green, K.; Kempka, D.; Lackey, L. 1994. Using remote sensing to detect and monitor land-cover and land-use change, *Photogrammetric Engineering and Remote Sensing* 60(3): 331–337.

Grünreich, D. 1985. Ein Vorschlag zum Aufbau einer großmaßstäbigen topographisch-kartographischen Datenbank unter besonderer Berücksichtigung der Grundrißdatei des ALK-Systems, *Nachrichten aus dem Karten- und vermessungswesen* 1:53–62.

Hame, T.; Heiler, I.; San Miguel-Ayanz, J. 1998. An unsupervised change detection and recognition system for forestry, *International Journal of Remote Sensing* 19(6): 1079–1099. <http://dx.doi.org/10.1080/014311698215612>

Hampe, M.; Sester M.; Harrie L. 2004. Multiple Representation Databases To Support Visualisation On Mobile Devices, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 35.

Harrie L.; Sarjakoski L. T.; Lehto L. 2002. A variable-scale map for small-display cartography, *Proceedings of the Joint International Symposium on “GeoSpatial Theory, Processing and Applications”*.

Harrie, L.; Hellstrom, A. K. 1999. A case study of propagating updates between cartographic data sets, *Proceedings/Actes, 19th International Cartographic Conference, 11th General Assembly of ICA*.

Hemmatnia, E.; van den Broecke, J.; van Raamsdonk, K. 2010. Harmonising Duch National Geodata Conformant to INSPIRE Using Combined Transformation, *International Journal of Spatial Data Infrastructure Research* 5: 365–381.

Herold, M.; Latham, J. S.; di Gregorio, A.; Schumilius, C. C. 2006a. Evolving standards in land cover characterization, *Journal of Land Use Science* 1: 2–4. <http://dx.doi.org/10.1080/17474230601079316>

Herold, M.; Woodcock, C.E.; di Gregorio, A.; Mayaux, P. 2006b. A Joint Initiative for Harmonization and Validation of Land Cover Datasets, *Geoscience and Remote Sensing* 44(7): 1719–1727. <http://dx.doi.org/10.1109/TGRS.2006.871219>.

Higer, C.; Klien, E.; Rix, J.; Zlatanova, S. 2013. Geospatial Data Harmonisation: Methods, Tools and Benefits, 200 p.

Hintz, D. 2012. Data Harmonization Principles and Development Approaches as Applied to INSPIRE SDIs. [interaktyvus] [žiūrėta 2013 m. rugsėjo 7 d.]. Prieiga per internetą:

<http://cdn.safe.com/resources/fme/Data-Harmonization-Principles-INSPIRE.pdf>

Hobona, G.; Jackson, M.; Gould, M.; Higgins, C.; Brauner, J.; Matheus, A.; James, P. 2009. Establishing a persistent interoperability testbed for European geospatial research, *AGILE 2009 Conference*. [interaktyvus] [žiūrėta 2013 m. rugsėjo 7 d.]. Prieiga per internetą:

http://www.agile-online.org/conference_paper/cds/agile_2009/agile_cd/pdfs/31.pdf

INSPIRE Data Specification on Coordinate Reference System – Technical Guidelines. D2.8.I.1_v3.2. [interaktyvus] [žiūrėta 2014 m. birželio 2 d.]. Prieiga per internetą: http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_RS_v3.2.pdf

INSPIRE Drafting Team "Data Specifications" Methodology for the development of data specifications. 2008. [interaktyvus] [žiūrėta 2014 m. birželio 2 d.]. Prieiga per internetą:

http://inspire.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.6_v3.0.pdf

INSPIRE Generic Conceptual Model D2.5_v3.4. 2014. [interaktyvus] [žiūrėta 2014 m. birželio 2 d.]. Prieiga per internetą:

http://inspire.ec.europa.eu/documents/Data_Specifications/D2.5_v3.4.pdf

Įsakymas dėl valstybinės koordinačių sistemos įvedimo tvarkos. 1996. [interaktyvus] [žiūrėta 2014 m. birželio 2 d.]. Prieiga per internetą: <https://www.e-tar.lt/portal/lt/legalAct/TAR.6D575923F94A>

International Cartographic Association. Commission II. 1973. *Multilingual Dictionary of Technical Terms in Cartography*, 573 p.

Jansen, L. J. M. 2005. Harmonization of land use class sets to facilitate compatibility and comparability of data across space and time, *Journal of Land Use Science* 1(2-4):127–156. 10.1080/17474230601079241

Jansen, L. J. M.; Groom, G. 2004. Thematic harmonisation and analyses of Nordic data sets into Land Cover Classification System (LCCS) terminology, *Developments in Strategic Landscape Monitoring for the Nordic Countries* 91–111.

Jansen, L. J. M.; Groom, G.; Carrai, G. 2008. Land-cover harmonisation and semantic similarity: some methodological issues, *Journal of Land Use Science* 3(2-3): 131–160. <http://dx.doi.org/10.1080/17474230802332076>

Jones, C. B.; Abdelmoty, A. I.; Loneragan, M. E.; van der Poorten, P.; Zhou, S. 2000. Multi-Scale Spatial Database Design for Online Generalisation, *Proceedings of the 9th International Symposium on Spatial Data Handling* 7b: 32–44.

Jones, C. B.; Bundy, G.L.; Ware, J. M. 1995. Map generalization with a triangulated data structure, *Cartography and Geographic Information Systems* 22(4): 317–331.

Keates, J. 1989. *Cartographic Design and Production. Second Edition*, Longman Scientific, 261 p.

Kiehle, C.; Greve, K.; Heier, C. 2007. Requirements for Next Generation Spatial Data Infrastructures Standardized Web Based Geoprocessing and Web Service Orchestration, *Transactions in GIS* 11(6): 819–834. <http://dx.doi.org/10.1111/j.1467-9671.2007.01076.x>

Kilpelainen, T. 1997. *Multiple Representation and Generalization of Geo-Databases for Topographic Maps*, Ph.D. thesis, Finnish Geodetic Institute, Helsinki University of Technology, 124 p.

Kilpelainen, T. 2000. Maintenance of multiple representation databases for topographic data, *The Cartographic Journal* 37(2): 101–107. 10.1179/0008704.37.2.p101

Kreveld, M. C. 2001. Smooth generalization for continuous zooming, *Proceedings of 20th international cartographic conferences*.

Lamy, S.; Ruas, A.; Demazeau, Y.; Jackson, M.; Mackaness, W.; Weibel, R. 1999. The application of Agents in Automated Map Generalisation, *Proceedings of 19th ICA meeting*.

Lecordix, F.; Jahard, Y.; Lemarié, C.; Hauboin, E. 2005. The end of Carto2001 Project TOP100 based on BD Carto® database, *8th ICA Workshop on Generalization and Multiple Representation*. [interaktyvus] [žiūrėta 2013 m. gruodžio 21 d.]. Prieiga per internetą:

http://generalisation.icaci.org/images/files/workshop/workshop2005/Lecordix_Jahard_et_al_.pdf.

Lehto, L. 2007a. *Real-Time Content Transformations in a Web Service-Based Delivery Architecture for Geographic Information*, PhD Thesis, Helsinki University of Technology, 65 p. [interaktyvus] [žiūrėta 2011 m. vasario 3 d.]. Prieiga per internetą: <http://lib.tkk.fi/Diss/2007/isbn9789517112727/isbn9789517112727.pdf>

Lehto, L. 2007b. Schema Translations in a Web Service Based SDI, *Proceedings 10th AGILE Conference on Geographic Information Science*. [interaktyvus] [žiūrėta 2011 m. vasario 3 d.]. Prieiga per internetą: http://people.plan.aau.dk/~enc/AGILE2007/PDF/29_PDF.pdf.

Lehto, L. 2009. Best Practice for Content Transformations Enabling INSPIRE-Compliant Data Delivery. Deliverable 11.1. Public report. ESDIN, European Spatial Data Infrastructure Network, ECP-2007-GEO-317008, 44 p.

Lemarie, C.; Raynal, L. 1996. Geographic data matching: first investigations for a generic tool, *GIS/LIS'96* 405–420.

Lenczowski, R. 1997. The military as users and producers of global spatial data, *Framework for the world. Cambridge (UK), Geoinformation International* 85–110.

Letho, L.; Sarjakovski, T. 2004. Schema Translations by XSLT for GML-Encoded Geospatial Data in Heterogeneous Web-Service Environment, *Proceedings XXth ISPRS Congress, Istanbul* 12–23.

Li, L.; Fei, L. 2008. Study on the change detection model and method for navigation digital map data, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science* 37(B2): 57–62. [interaktyvus] [žiūrėta 2012 m. sausio 23 d.]. Prieiga per internetą: http://www.isprs.org/proceedings/XXXVII/congress/2_pdf/1_WG-II-1/10.pdf

Li, L.; Goodchild, M. F. 2010. Optimized feature matching in conflation, *Geographic Information Science: 6th International Conference, GIScience* 14–17.

Li, L.; Goodchild, M. F. 2011. An optimisation model for linear feature matching in geographical data conflation, *International Journal of Image and Data Fusion* 2(4): 309–328. <http://dx.doi.org/10.1080/19479832.2011.577458>

Li, Z. L.; Ho, A. 2004. Design of Multi-Scale and Dynamic Maps for Land Vehicle Navigation, *The Cartographic Journal* 41(3): 265–270.

Li, Z. L.; Su, B. 1996. Algebraic models for feature displacement in the generalization of digital map data using morphological techniques, *Cartographica* 32(3): 39–56.

- Li, Z. L.; Sui, H. G. 2000. An integrated technique for automated generalisation of contour maps, *The Cartographic Journal* 37(1): 29–37.
- Li, Z. L.; Yan, H.; Ai, T.; Chen, J. 2004. Automated Building Generalization Based on Urban Morphology and Gestalt Theory, *Journal of Geographical Information Science* 18(5): 513–534.
- Li, Z.; Wong, M. 2008. Animating basic operations for digital map generalization with morphing techniques, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 37(B2): 637–642.
- Lichtner, W. 1979. Computer-Assisted Processing of Cartographic Generalization in Topographic Maps, *Geo-Processing*, 1:183–199.
- Lu, D.; Mausel, P.; Brondizio, E.; Moran, E. 2004. Change detection techniques, *International journal of remote sensing* 25(12): 2365–2401.
- Luccy, R.; Millot, M. 2009. INSPIRE Invoke Services. Survey on requirements, challenges and recent research findings supporting the development of the Invoke spatial data service specification, 27 p. [interaktyvus] [žiūrėta 2012 m. kovo 23 d.]. Prieiga per internetą:
http://inspire.ec.europa.eu/documents/Network_Services/Technical_notes_inspire_invoke_services.pdf
- Lunetta, R. S.; Knight, J. F.; Ediriwickrema, J.; Lyon, J. G.; Worthy, L. D. 2006. Land-cover change detection using multi-temporal MODIS NDVI data, *Remote sensing of environment* 105(2): 142–154. <http://dx.doi.org/10.1109/amtrsi.2005.1469870>
- Mackaness W. A.; Mackechnie G. 1999. Detection and Simplification of Road Junctions in Automated Map Generalization, *GeoInformatica* 3: 185–200.
- Mas, J. F. 1999. Monitoring land-cover changes: a comparison of change detection techniques, *International journal of remote sensing* 20(1): 139–152. <http://dx.doi.org/10.1080/014311699213659>
- Masser, I. 2005. *GIS Worlds: Creating Spatial Data Infrastructures*, Redlands, California, ESRI Press, 312 p.
- McConneli, W. J.; Moran, E. F. 2001. Meeting in the middle: the challenge of meso-level integration. An international workshop on the harmonization of land-use and land-cover classification. Lucc Report Series No. 5. Anthropological Center for Training and Research on Global Environmental Change - Indiana University and Lucc International Project Office, Louvain-la-Neuve.
- McConnell, W. J.; E. F. Moran. 2001. Meeting in the middle: the challenge of meso-level integration. An international workshop on the harmonization of land-use and land-cover classification. Lucc Report Series No. 5.
- McHaffie, P. H. 2002. Towards the automated map factory: early automation at the U.S. Geological Survey, *Cartography and Geographic Information Science* 29(3): 193–206. <http://dx.doi.org/10.1559/152304002782008521>

- McMaster R. B.; Shea K. S. 1988. Cartographic generalization in a Digital Environment: a framework for implementation in a geographic information system, *GIS/LIS'88 Proceedings* 240–249.
- McMaster, R. 1989. The Integration of Simplification and Smoothing Algorithms in Line Generalization, *Cartographica* 26 (1):101–121.
- McMaster, R. B; Shea, K. S. 1992. *Generalization in Digital Cartography*, Association of American Geographers, 134 p.
- McMaster, R; Monmonior, M. 1989. A conceptual framework for quantitative and qualitative raster-mode generalisation, *Proceedings of GIS/LIS'89* 390–403.
- Milieškaite, J. 2010. Kovariacijos metodo taikymas analizuojant (palyginant) skaitmeninius vaizdus kartografijs, *Geodezija ir Kartografija* 36(1): 30-34. <http://dx.doi.org/10.3846/gc.2010.05>
- Milieškaitė, J. 2011. The application of the covariance method analysing the digital images of land surface, *Geodesy and Cartography* 37(3): 105-110. <http://dx.doi.org/10.3846/13921541.2011.626260>
- Milieškaitė, J., & Vaitkus, G. (2011). LANDSAT TM kosminių nuotraukų segmentavimo metodų palyginimas atliekant žemės dangos klasifikavimą, *Geodesy and Cartography* 37(3): 135-142.
- Monmonior, M. 1983). Raster-mode area generalization of land use and land cover maps, *Cartographica* 20(4): 64–91.
- Müller, J. C. 1991. Generalisation of Spatial Databases, *Geographical Information Systems* 1: 457–475.
- Neumann, K.; Herold, M.; Hartley, A.; Schmullius, C. 2007. Comparative assessment of CORINE2000 and GLC2000: Spatial analysis of land cover data for Europe, *International Journal of Applied Earth Observation and Geoinformation* 9(4): 425–437. <http://dx.doi.org/10.1016/j.jag.2007.02.004>
- Nickerson, B. 1988. Automated cartographic generalization for line features, *Cartographica* 25(3): 15–66.
- Nickerson, B.G.; Freeman H. 1986. Development of a Rule – Based Systems for Automatic Map Generalization, *Proceedings 2nd Intern Symp on Spatial Data Handling*.
- Pammer, A.; Hopfstock, A.; Ipsa, A.; Vanova, J.; Vilius, I.; Delattre, N. 2010. EuroRegionalMap – How to Succeed in Overcoming National Borders, *Cartography in Central and Eastern Europe, Lecture Notes in Geoinformation and Cartography* 19–40. http://dx.doi.org/10.1007/978-3-642-03294-3_2.
- Papšienė, L.; Budrevičius, J. D.; Marma, M. 2013a. Lietuvos Respublikos teritorijos M1:50 000 georeferencinių erdvinių duomenų rinkinio atnaujinimo galimybių tyrimo, naudojant automatizuotus generalizavimo metodus ataskaita 147 p.

Peerbocus, A.; Jomier, G.; Badard, T. 2002. A methodology for updating geographic databases using map versions, *Advances in Spatial Data Handling* 363–376.

Plazanet, C.; Affholder, J. G.; Fritsch, E. 1995. The importance of geometric modeling in linear feature generalization, *Cartography and Geographic Information Systems* 22(4): 291–305.

Podolskaya, E.S.; Anders, K. H.; Haunert J. H.; Serester, M. 2007. Quality Assessment for Polygon Generalization, *5th International Symposium on Spatial Data Quality*. [interaktyvus] [žiūrėta 2011 m. spalio 5 d.]. Prieiga per internetą: http://www.isprs.org/proceedings/XXXVI/2-C43/Session3/paper_Podolskaya_Anders_Haunert_Sester.pdf

Radke, R. J.; Andra, S.; Al-Kofahi, O.; Roysam, B. 2005. Image change detection algorithms: a systematic survey, *IEEE Transaction on Image Processing* 14(3), 294–307. <http://dx.doi.org/10.1109/tip.2004.838698>

Ratajski L. 1976. Cartology: its developed concept, *The Polish Cartographer* 7–23.

Refsgaard, J. C.; Nilsson, B.; Brown, J.; Klausner, B.; Moore, R.; Bech, T.; Vurro, M.; Blind, M.; Castilla, G.; Tsanis, I.; Biza, P. 2005. Harmonised techniques and representative river basin data for assessment and use of uncertainty information in integrated water management (HarmoniRiB), *Environmental Science & Policy* 8(3): 267–277. <http://dx.doi.org/10.1016/j.envsci.2005.02.001>

Reitz, T.; Kuijper, A. 2009. Applying Instance Visualisation and Conceptual Schema Mapping for Geodata Harmonisation, *Lecture Notes in Geoinformation and Cartography* 173–194. http://dx.doi.org/10.1007/978-3-642-00318-9_9

Reitz, T.; Templer, S. 2012. An Environment for the Conceptual Harmonisation of Geospatial Schemas and Data. [interaktyvus] [žiūrėta 2013 m. balandžio 2 d.]. Prieiga per internetą: http://www.agile-online.org/Conference_Paper/CDs/agile_2012/proceedings/papers/Paper_Reitz_An_Environment_for_the_Conceptual_Harmonisation_of_Geospatial_Schemas_and_Data_2012.pdf

Rhind, D. W.; Clark, P. 1988. Cartographic inputs to global databases. In Mounsey H M, Tomlinson R F (eds) *Building databases for global science*. London, Taylor and Francis: 79–104.

Richardson, D.E. 1993. *Automatic Spatial and Thematic Generalization using a Context Transformation Model*. Ph.D thesis, Wageningen Agricultural University, The Netherlands.

RISE. 2007. *Methodology & Guidelines on Use Case & Schema Development*. Version 1.2. [interaktyvus] [žiūrėta 2011 m. spalio 5 d.]. Prieiga per internetą: <http://www.eurogeographics.org/documents/RISE15MethodologyandGuidelinesV1.2.pdf>

Robinson, A.H.; Sale, R.; Morrison, J.L.; Muehrcke, P.C. 1984. *Elements of Cartography*. (5th Edition), Wiley, 688 p.

- Ruas, A. 1999. *Modèle de généralisation de données géographiques à base de contraintes et d'autonomie*, Ph.D. thesis Université de Marne la Vallée, Paris, 323 p.
- Ruas, A.; Lagrange, J.P. 1995. Data and knowledge modelling for generalisation, *GIS and generalization, Methodology and Practice, GISDATA 1*: 73–90.
- Ruas, A.; Plazanet, C. 1996. Strategies for automated generalization, *Proceedings of the 7th International Symposium on Spatial Data Handling, Advances in GIS Research* 2(6): 1–18.
- Rusak M. E.; Castner H. W. 1990. Horton's ordering scheme and the generalisation of river networks, *The Cartographic Journal* 27: 104–112.
- Safra, E.; Kanza, Y.; Sagiv, Y.; Beerli, C.; Doytsher, Y. 2010. Location-based algorithms for finding sets of corresponding objects over several geo-spatial data sets, *International Journal of Geographical Information Science* 24(1): 69–106.
- Salichtchev, K. A. 1976. History and contemporary development of cartographic generalization, *International Yearbook of Cartography* 16:158–72.
- Sandgren, U.; Ottoson, C. 2002. Harmonisation of European Road Data: How to Create Large Scale Interoperability Between National Databases, *Proceedings of the 8th European Comission GI&GIS Workshop: ESDI: A Work in Progress*. http://www.ec-gis.org/Workshops/8ec-gis/cd/papers/3_crd_us.pdf
- Savino, S. 2011. *A solution to the problem of the generalization of the Italian geographical databases from large to medium scale: approach definition, process design and operators implementation*, PhD Thesis. Università Di Padova. 139 p.
- Shea K. S.; McMaster R. B. 1989. Cartographic generalization in a Digital Environment: when and how to generalize, *AUTO-CARTO, 9th International Symposium on Computer Assisted Cartography* 56–67.
- Shi, W.; Cheung, C. 2006. Performance Evaluation of Line Simplification Algorithms for Vector Generalization, *The Cartographic Journal* 43(1): 27–44.
- Simav, O.; Aslan, S.; Cetinkaya, B.; Cobankaya, O. N. 2010. Implementation of Comprehensive Modeling Techniques on KARTOGEN Generalisation Software, *13th ICA Workshop on Generalisation and Multiple Representation*. [interaktyvus] [žiūrėta 2013 m. gruodžio 21 d.]. Prieiga per internetą: http://www.hgk.msb.gov.tr/haritalar_projeler/bildiriler/kartografya/makale%28pdf%29/karto_tek_bil9.pdf
- Singh, A. 1989. Review Article Digital change detection techniques using remotely-sensed data, *International Journal of Remote Sensing* 10(6): 989–1003.
- Smith, N.S.; Rhind, D. W. 1999. Characteristics and source of framework data, *Geographic information system* 2:655–666. http://www.geos.ed.ac.uk/~gisteac/gis_book_abridged/files/ch47.pdf
- Smits, P. C.; Friis-Christensen, A. 2007. Resource Discovery in a European Spatial Data Infrastructure, *IEEE Transactions on Knowledge and Data Engineering* 19(1): 85–95.

Spaccapietra, S.; Parent, C. 2000. GIS databases: From multiscale to multirepresentation, *SARA 2000* 57–70.

Steiniger, S.; Hunter, A. J. 2012. Free and open source GIS software for building a spatial data infrastructure, *Lecture Notes in Geoinformation and Cartography* 247–261. http://dx.doi.org/10.1007/978-3-642-10595-1_15

Steiniger, S.; Meier, S. 2004. Snakes: a technique for line smoothing and displacement in map generalisation, *7th ICA workshop on generalisation and multiple representation*.

Stoter, J. 2010. *State-of-the-Art of Automated Generalisation in Commercial Software, Report of EuroSDR projects. Gopher*, Amsterdam, 265 p.

Stoter, J.; Nijhuis, R.; Post, M.; van Altena, V.; Bulder, J.; Bruns, B.; van Smaalen, J. 2011. *Feasibility study on an automated generalisation production line for multiscale topographic products*.

Stoter, J.; Quak, W.; Hofman, A. 2009. Harmonising and Integrating Two Domain Models Topography, SDI Convergence – Resach, Emerging Trends, and Critical Assesment 89–106.

Stoter, J.; Van Oosterom, P.; Ploeger, J.; Hendrik, D.; Aalders, H. 2004. Conceptual 3D Cadastral Model Applied in Several Countries in TS25 – Appropriate Technologies for Good Land Administration II – 3D Cadastre, *FIG Working Week*.

Su, B.; Li, Z. L. 1995. An algebraic basis for digital generalisation of area-patches based on morphological techniques, *The Cartographic Journal*, 32(2): 148–153.

Su, B.; Li, Z. L.; Lodwick, G. 1998. Algebraic models for collapse operation in digital map generalization using morphological operators, *Geoinformatica*, 2(4): 359–382.

Su, B.; Li, Z. L.; Lodwick, G., 1997a. Morphological transformation for the elimination of area features in digital map generalization, *Cartography* 26(2): 23–30.

Su, B.; Li, Z. L.; Lodwick, G.; Müller, J. C. 1997b. Algebraic Models for the Aggregation of Area Features Based upon Morphological Operators, *International Journal of Geographical Information Science* 11(3): 233–246.

Technical Guidance for implementation of INSPIRE Download Service. Version 3.1. 2013a.

http://inspire.ec.europa.eu/documents/Network_Services/Technical_Guidance_Download_Services_v3.1.pdf

Technical Guidance for implementation of INSPIRE View Service. Version 3.11. 2013b. http://inspire.ec.europa.eu/documents/Network_Services/TechnicalGuidance_ViewServices_v3.11.pdf

Technical Guidance for INSPIRE Schema Transformation Network Service. Version 3.0. 2010. http://inspire.ec.europa.eu/documents/Network_Services/JRC_INSPIRE-TransformService_TG_v3-0.pdf

- Thompson R. C.; Richardson D. E. 1995. A Graph Theory Approach to Road Network Generalization, *Proceedings of the 17th ICA Meeting* 1871–1880.
- Thompson R. C.; Richardson D. E. 1999. The 'Good Continuity' Principle of Perceptual Organisation applied to the Generalisation of Road Networks, *Proceeding of the 19th International Cartographic Conference* 1215–1225.
- Timpf, S. 1998. Map Cube Model – a model for multi-scale data, *Proceedings 8th International Symposium on Spatial Data Handling (SDH'98)* 190–201.
- Töpfer F.; Pillewizer W. 1966. The principles of selection, *The Cartographic Journal* 3(1): 10–16.
- Torun, A.; Köbben B.; Lemmens R. 2000. Processing Spatial Data on the Internet, *International Archives of Photogrammetry and Remote Sensing* 33(B6): 269–278.
- Tryfona, N.; Max, J. 1996. Multi-Resolution Spatial Databases: Consistency Among Networks, *6th International Workshop on Foundations of Models and Languages for Data and Objects: Integrity in Databases*.
- Ulugtekin, N.N.; Dogru A.O.; Thomson R.C. 2004. Modeling Urban Road Networks Integrating Multiple Representations of Complex Road and Junction Structures, *Proceedings of the 12th International Conference on Geoinformatics* 757–764.
- Urbanas, S. 2001a. *Linijinių kartografinių objektų generalizacijos metodai objektiškai orientuotose geoinformacinėse sistemose*, Daktaro disertacija, p.111.
- Urbanas, S. 2001b. Objektiškai orientuotos duomenų bazės kartografinės generalizacijos procese, *Geodezija ir kartografija* 27(2): 68–73.
<http://dx.doi.org/10.1080/13921541.2001.10552915>
- Villa, P.; Reitz, T.; Gomasasca, M. A. 2008. The HUMBOLDT project for data harmonisation in the framework of GMES and ESDI: Introduction and early achievements. International Society for Photogrammetry and Remote Sensing - Proceedings of Commission IV. S. 1741–1746. Beijing, China
http://www.isprs.org/proceedings/XXXVII/congress/4_pdf/303.pdf
- Visvalingham, M.; Whyatt, J. 1993. Line generalization by repeated elimination of points, *The Cartographic Journal* 30(1): 46–51.
- Volpi, M.; Tuia, D.; Bovolo, F.; Kanevski, M.; Bruzzone, L. 2013. Supervised change detection in VHR images using contextual information and support vector machines. *International Journal of Applied Earth Observation and Geoinformation* 20: 77–85.
- Walter, V.; Fritsch, D. 1999. Matching spatial data sets: a statistical approach. *International Journal of Geographical Information Science* 13(5): 445–473.
- Walters, R.; Beare, M.; Walker, R.; Milot, M. 2011. Schema Transformation for INSPIRE. IJSDIR. 6.
- Wang, Z.; Müller J. C. 1998. Line Generalisation Based on Analysis of Shape Characteristics, *Cartography and Geographic Information Systems* 25(1): 3–15.

Web Feature Service (WFS) Implementation Standart V1.1.0, OGCR Implementation Standart. Open Geospatial Consortium Inc., 2005.

Web Map Service (WMS) Implementation Standart V1.3.0, OGCR Implementation Standart. Open Geospatial Consortium Inc., 2006.

Weibel R. 1991. Amplified intelligence and rule-based systems, *Map generalization: making rules for knowledge representation* 172–186.

Weibel, R.; Dutton, G. 1999. Generalising spatial data and dealing with multiple representations, *Geographical information systems* 1: 125–155.

West-Nielsen, P.; Meyer, M. 2007. Automated generalisation in a map production environment – the KMS experience, *Generalisation of geographic information: Cartographic modelling and applications* 301–314. <http://dx.doi.org/10.1016/b978-008045374-3/50018-1>

Witschas, S. 2005. Cross-border mapping – experience with geodata and geonames, *XXII International Cartographic Conference*. http://www2.ioer.de/recherche/pdf/2005_witschas_icc.pdf

Yuhong. W.; Qiang, W.; Haitao, J. 2013. A Schema Matching Approach Based on Spatial Instances for Propagating Updates Between GIS Databases, *International journal on Advances in Information Sciences and Service Sciences* 5(8): 171–179. <http://dx.doi.org/10.4156/aiss.vol5.issue8.21>

Zhu, Y.; Zhou, S.; Lu, T. 2011. Research on Spatial Data Line Generalization Algorithm in Map Generalization, *Journal of Software* 6(2): 241–248.

Zlatanova, S.; Stoter J.E.; Quak W. 2004. Management of multiple representations in spatial DBMSs, *Proceedings of the 7th AGILE Conference on Geographic Information Science* 269–278.

Autorės mokslinių publikacijų disertacijos tema sąrašas

Straipsniai recenzuojamuose mokslo žurnaluose

Papšienė, L. (Папшене, Л.). 2013a. Автоматизированная генерализация в процессе обновления базовых пространственных данных мелкого масштаба. *Інженерна геодезія: Науково-технічний збірник* 60: 116–128.

Papšienė, L.; Papšys, K.; Vaitkevičienė, J. 2013b. A model to detect cartographic base changes affecting the map update, *Aviation* 17(1): 22–29.

Papšienė, L.; Papšys, K. 2012 Changes affecting generalization of land cover features in a smaller scale, *Geodesy and cartography*, 38(3): 98–105.

Papšienė, L.; Papšys, K. 2011. Possibilities of updating small-scale basic spatial data in Lithuania using generalization methods, *Geodesy and cartography* 37(4): 143–148.

Beconytė, G.; Kryžanauskas, A.; Papšienė, L. 2010a. Lietuvos erdvinės informacijos sklaidos galimybės ir perspektyvos, *Geodezija ir kartografija* 36(2): 73–80.

Straipsniai kituose leidiniuose

Papšienė, L.; Budrevičius, J. B.; Marma, M. 2014. The Automated Update of Cartographic Data at a Scale of 1:50,000 in Lithuania: Problems and Solutions, *in Proceedings of the 8th International Conference "Environmental Engineering", 2014, Vilnius, Lithuania: selected papers.*

Papšienė, L. 2013b. Challenges of Automated Processes of Update of Reference Spatial Data at a Scale of 1:50.000 in Lithuania, *in Proceedings of SGEM2013 Conference, June 16–22, 2013. 1: 661-668. ISSN 1314-2704.*

Beconytė, G.; Kryžanauskas, A.; Papšienė, L. 2010b. Lietuvos erdvinės informacijos infrastruktūros plėtros perspektyvos, *6-osios mokslinės konferencijos „Mokslas Gamtos mokslų fakultete“, įvykusios Vilniuje 2010 m. spalio 22 d. pranešimų medžiaga.* Vilnius: Vilniaus universiteto leidykla, 125–134. ISBN 978-9955-33-602-0.

Summary in English

Introduction

Problem's formulation

In order to represent and analyse natural and anthropogenic phenomena of the real world, they had for centuries been interpreted, mapped and represented in paper maps. Meanwhile, appearance of geographic information systems (GIS) offered better ways of storing and updating spatial data than old style paper maps, so the mentioned objects are stored as spatial data nowadays with possibilities to join them as separate themes of spatial data. Therefore now, paper maps are just one of many ways of cartographic representation of such information. Development of GIS technologies allowed storing huge volumes of spatial data of various accuracy describing both quantitative and qualitative attributes of objects.

Depending on the theme and accuracy, spatial data are created or updated by various means. For example, for the creation of (geo)referential or topographic spatial data the following methods may be applied: surveying, photogrammetry methods, generalisation of spatial data of a larger scale, etc. Unfortunately, quite often similar spatial data sets of different scales are updated separately and independently from each other applying different data models and standards; besides, data are stored in different data formats and updated at different times. This causes problems of compatibility and timely update of spatial data sets. In particular, it becomes important when attempting to harmonise spatial data both at the same and different scales from different institutions or

even different states. Non-harmonised spatial data sets impede the use and analysis of several sets concurrently.

Harmonisation of spatial data is a process of harmonising information of different but related spatial data sets. In the last decade, harmonisation procedures has mainly focused on restructuring and transformation of spatial data by harmonising spatial data sets of the same or similar scales and of the same theme, while compatibility of spatial data of different scales has been left apart. Therefore, compatibility of spatial or cartographic data among spatial data sets of different scales requires the improvement of harmonisation methodology for spatial data, which would, in addition, cover generalisation possibilities of spatial data for changing the type of elements of spatial data, decreasing their volume or simplifying their complexity by excluding insignificant, but retaining significant attributes of the represented objects.

Relevance of the thesis

Harmonisation of different spatial or cartographic data sets requires a lot of human resources. This is why identical spatial data in different sets are duplicated and rarely updated. Automation of the harmonisation process in relating spatial data sets of different scales would allow applying the principle that spatial data would be captured once and at the most appropriate scale (Directive of the European Parliament and of the Council ... 2007), and would give a possibility for a user to receive smaller scale data harmonised with spatial data sets at a larger scale. However, depending on the volume of harmonised data the automation process may require considerable technical resources. For this reason it is also essential to take into account the possibility to harmonise only the changed data.

Furthermore, spatial data infrastructures (SDIs) cover the technological means designed to provide copies of spatial data sets, their parts and interactive maps based on spatial data for users, also the means, which enable harmonisation of different spatial data sets and their presentation as a whole. So, SDIs should include harmonisation models of spatial data covering a whole of technological means for harmonisation of spatial data.

Research object

Methodology for harmonisation of spatial data as well as modelling and generalisation methods of spatial data applied in digital cartography. The object of the experiment is (geo)referential spatial data sets of the Republic of Lithuania at scales M 1:10,000 (GDR10LT) and M 1:50,000 (GDR50LT).

Aim of the thesis

The aim of the thesis is to improve harmonisation methodology for spatial data and data modelling procedure, when data have to be transformed into a spatial data set at a smaller scale taking into account the significance of changes in the data source.

Objectives of the thesis

To achieve the aim of the thesis, the following objectives were formulated:

1. To evaluate algorithms for simplification and generalisation of line and polygon elements of spatial data to be applied in harmonisation of spatial data representing natural and anthropogenic objects;
2. To analyse the influence of spatial data changes in the source data set on the transformation spatial data in the harmonisation process;
3. To improve methodology for identification and evaluation of changes in polygon spatial data applied in digital cartography;
4. To improve harmonisation methodology for spatial data, which would allow developing procedures to be applied in SDIs taking into account generalisation methods of spatial data applicable in digital cartography.

Research methodology

The thesis used foreign and national scientific publications and researches for reference as well as theoretical and experimental research on selection, simplification and generalisation algorithms for spatial data elements applying GIS technologies and modelling a change in elements of identical spatial data in source spatial data sets; a comparative analysis of research results was carried out.

Scientific novelty of the thesis

The dissertation produced the following results new to the measurement engineering:

1. Improved methodology for identification and evaluation of influence for a transformation process of changes in elements of spatial data, which are applicable in harmonisation of spatial data when relating spatial data sets at the identical and different cartographic scales.
2. Improved harmonisation model for polygon elements of spatial data taking into account the topological relations and different tolerances and algorithms applied.
3. Improved methodology for harmonisation of spatial data in SDIs environment

Practical value of research findings

The improved harmonisation methodology may be applied in SDIs environment for implementation of models which can be used for harmonisation of spatial data between GDR10LT and GDR50LT data sets, also it is possible to use this methodology for harmonisation between other large and small cartographic scale spatial datasets.

Defended statements

1. Suggested harmonisation methodology for spatial data is suitable for the harmonisation of spatial data sets at different scales.

2. During process of the developing a harmonisation model for the spatial data of related sets at different cartographic scales, it is necessary to set up a priority sequence for cartographic objects.
3. In the harmonisation process, potential places to be updated in spatial data sets are identified with the help of the proposed identification and evaluation methodology for spatial data changes.
4. Proposed harmonisation procedures for spatial data may detect and eliminate insignificant changes of spatial data elements.

Approval of research findings

The subject of the dissertation was presented in eight scientific articles: one was published in periodicals referenced in international data bases (Papšienė 2013a); two in publications of international conferences (Papšienė *et al.* 2014, Papšienė 2013b); four in other approved data bases (Papšienė *et al.* 2012; Papšienė, Papšys 2012, 2011, Beconytė *et al.* 2010); and one in the material of a national conference (Beconytė *et al.* 2010).

The research findings were presented at nine scientific conferences and seminars in Lithuania and abroad:

- 9th International Conference “*Environmental Engineering*” in Vilnius, 2014;
- 13th International Multidisciplinary Scientific GeoConference SGEM, in Albena, 2013;
- Scientific Conference “*From Cartography to Science of Geographic Information*” in Vilnius, 2013;
- 15th Conference of Young Lithuanian Scientists “*Science – the future of Lithuania*” in Vilnius, 2012;
- RTU 53rd International Scientific Conference in Riga, 2012;
- Scientific Conference “*Aviacija 2012*” in Vilnius, 2012;
- 14th Conference of Young Lithuanian Scientists “*Science – the future of Lithuania*” in Vilnius, 2011;
- Scientific conference “*Civil Engineering and Geodesy*” in Vilnius, 2010;
- Scientific conference “*Science in the Faculty of Natural Science*” in Vilnius, 2010.

Structure of the thesis

The thesis comprises an introduction, three chapters, general conclusions, the list of references, the list of published scientific works on the topic of dissertation, summary in English and annexes.

The volume of the thesis amounts to 162 pages with the exception of annexes; the text contains 21 numbered equations, 40 illustrations and 30 tables. The dissertation is built upon 187 references.

Acknowledgements

I would like to express my very sincere appreciation to my scientific supervisor Professor Dr Eimuntas Kazimieras Paršeliūnas for scientific advice, valuable guidance,

optimism, patience and assistance in the preparation of the dissertation. I also thank Head of Geodesy and Cartography Department of Vilnius Gediminas Technical University Professor Dr Česlovas Aksamitauskas, department staff, Dr Giedrė Beconytė from the Centre of Cartography of Vilnius University, Director of SE The National Center of Remote Sensing and Geoinformatics „GIS-Centras“ Evaldas Rožanskas and colleagues for all the valuable suggestions and assistance. My grateful thanks are also extended to the staff of the Surveying, Mapping and Spatial Information Department of the National Land Service under the Ministry of Agriculture for the possibility to access the spatial data and study results. Especially I want to thank my family for their support and assistance during all my study period.

1. Analysis of harmonisation problems of spatial data in digital cartography

GIS technologies allowed effectively collecting and processing high volumes of spatial data on various themes. However, a large amount of non-harmonised spatial data sets do not provide broad opportunities for their use. Therefore, harmonisation of spatial data became a major challenge in unification of different data sets. Literary analysis has showed that definition of data harmonisation is ambiguous. In part, harmonisation is related to data standardisation. Masser (2005) by defining SDI also refers to standardisation of spatial data, which should be one of the major factors of SDI installation, since sharing of non-standardised data may lose its sense, as users may find it rather difficult to use non-standardised data. However, according to a definition, standardisation is the use of a uniform standard, which describes a specific data model and classifiers to define attributes, in collecting spatial data. Meanwhile, harmonisation is the unification of data, which are collected using different data models or standards (MacConnell, Morgan 2001). In the context of global use of spatial information, harmonisation is regarded as a possibility to harmonise various spatial data by integrating them into a single and unambiguous information data set satisfying the needs of a user (de Vries *et al.* 2007a). On 14 March 2007, Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) was adopted (INSPIRE... 2007). One of the main tasks of the INSPIRE programme is to enable the quality, interoperability and, where practicable, harmonisation of spatial data sets within Europe (INSPIRE Generic Conceptual Model... 2014). In the framework of SDI, interoperability is the possibility for spatial data sets to be combined and harmonised without repetitive manual intervention, in such a way that the result is coherent (INSPIRE... 2007, Law on Geodesy and Cartography 2010). In order to achieve this, data harmonisation may additionally be carried out, which would give access to spatial data via network services enabling to coherently integrate spatial data with other harmonised data with the help of a common uniform (standardised) data model. Thus, in general, harmonisation could be referred to as a process giving access to spatial data sets and transforming data by using a uniform data model so that users could use the entirety of interactive data regardless of the specific character of the initial data source.

There are basic cases of spatial data incompatibility, which affect modelling peculiarities of the harmonisation process (de Vries *et al.* 2007a, 2007b): data formats, spatial reference systems, conceptual schemas/data models, classifications schemes, scales/resolutions, levels-of-detail of data content, metadata, terminology, portrayal, multiple representation of the same spatial objects, spatial consistency at borders and processing functions.

Issues of transformation in the harmonisation process involve generalisation methods. In general, generalisation is a process covering the reduction of volume of spatial data and their complexity and the adaptation to a specific scale (Muller, 1991; Weibel, Dutton, 1999). This process generalises information and reduces its volume retaining specific significant cartographic elements of a map. The use of automatic procedures in generalisation is affected by three principal factors: reduction in work and time resources, the qualification and subjectivity of specialists (Kilpelainen 2000) and data accuracy achieved by manual update (McHaffie 2002). With the development of internet technologies, spatial data generalization is more important than ever. This has especially been noticeable in recent years because of the growing popularity of e-services providing spatial data sets for users. For generalisation on the web the researchers propose to use web based SDI and open source based OGC e-services. On the other hand, using only OGC e-services and invoke WPS e-services was a problematic task, since there was no mechanism for invoking the chain of e-services. Nevertheless, such type of e-services is necessary due to the complexity of generalisation algorithms, which include a set of various operations for generalisation. So, there were limitations for sophisticated mechanisms where models can call each other (Foster, Stoter 2006; Burghardt *et al.* 2005; Neun *et al.* 2008). During the INSPIRE implementation the new e-service is specified, which allows activating other e-services (INSPIRE Invoke Service) (European Parliament and of the Council of ... 2007, Lucci, Millot 2009).

Modern GIS technology allows extensive using of spatial data, integrating this data with other spatial data sets, analysing and developing them according to the needs. The process of determining changes in spatial data is an important part of a data harmonisation process. It allows reducing the use of resources for the update of spatial data. Besides, it enables preserving information about the unchanged elements of spatial data and tracking their life cycle. Most of the proposed ways how to automatically detect changes among different spatial data sets are based on the comparison of all spatial objects in spatial data sets before and after the update (Badart 1998) using the object-alignment algorithms (Lemarie *et al.* 1996). At the time of object alignment the relationships between spatial objects in different spatial data sets are determined. When such a relationship between the spatial data is determined changes can be analysed and classified according to their nature (Bahard 1998). Most of the algorithms designed to identify changes are for analysis of similar cartographic scale spatial datasets, so they are not suitable to detect changes in the larger-scale spatial data sets in order to assess their impact on the smaller scale spatial data sets.

Regarding the results of literature analysis and the aim of the thesis, the main tasks of the thesis are:

1. To select, research and evaluate algorithms for simplification and aggregation of linear and polygon features. The selected algorithms will be used in the harmonisation process during transformation of spatial data representing natural and anthropogenic objects.
2. To research possible changes in features of source spatial data and the relation between changes in features of source data and features of transformed and generalised spatial data.
3. To improve method for detection and evaluation of changes in features for digital cartography.

To improve methodology for harmonisation of spatial data, which can be used for the development of procedures, applied in SDI involving generalisation methods and to investigate practise of methodology.

2. The improvement of methodology of harmonisation of spatial data

The task of the thesis is to improve the methodology for the harmonisation of spatial data with a possibility to develop procedures, which will allow developing procedures for SDI and technologies, taking into account the possibility to harmonise spatial data sets at different cartographic scales using generalisation of spatial data.

The methodology was enhanced taking into account the general methodologies for spatial data harmonisation in SDI developed during the RISE (RISE in 2007) and HUMBOLD (Fichtinger *et al.* 2011) projects as well as regarding harmonisation components described in INSPIRE (INSPIRE Generic... 2014). The methodology revisited by the author of the thesis was improved in the light of the need to harmonise spatial datasets at different cartographic scales.

Development of spatial data harmonisation technology in most of the cases should include the same phases like in other IT projects: analysis, design, development, testing, and documentation.

The analysis phase is one of the most important stages. It includes the analysis of source and final spatial data sets, spatial data models, links between spatial data features and attributes for the identification of cardinality type between them (e.g. 1:1, 1>N, N:1, etc.), and topological relations (e.g., a feature must be covered by another feature or features must not have gaps between each other).

Design of technology for spatial data harmonization should include at least the following processes:

1. Accessibility and transferability of the primary spatial data set.
2. Automated validation of spatial data sources.
3. Transfer of spatial data to the temporal spatial database.
4. Transformation of spatial data. This process may consist of spatial data selection, renaming, re-classification, aggregation or disaggregation, generalisation, restructuration and addition of missing features.
5. Transfer of transformed spatial data to the final database.
6. Final verification of final spatial data.

Verification of source data will ensure the process of transformation, whereas the changed structure of the source spatial data or incorrectness of source features can influence the transformation process: it may not be available or an incorrect result may be obtained. Verification of the final spatial data evaluates the correctness of the transformation process. A negative result means that the source data might have changed and it has affected the result of the transformation.

During the transformation of spatial data, generalisation should be applied when:

1. Semantic consistency of spatial data is required or set in accrual rules of spatial data features. In this case, feature generalization using the same type of neighbouring cell aggregation algorithms should be used.
2. Identical spatial objects in source and final spatial data models are represented differently.
3. Cartographic scales of source and final spatial data differ.

In the development stage, spatial data harmonisation technology must be developed based on the results of a design stage and using the selected software and available IT. Testing, deployment and documentation are also applicable, as set out in typical IT development methodologies.

When developing spatial data generalisation in the transformation process, it is recommended to carry out tests concurrently. The tests should start with small areas of spatial data and subsequently increase them depending on the results.

The author presents some algorithms for some specific issues of generalisation considering the possibilities to use generalisation of spatial data in the transformation process. The thesis proposes algorithms for selection, elimination and aggregation of features, elimination of dangles of line features, polygon feature transformation to centre line and matching of boundaries of neighbouring line or polygon features.

The methodology for harmonisation of spatial data proposes several methods for detection and evaluation of changes in the source spatial data set as well:

- The method for identification of changes in spatial data using the Cyclic Redundancy Check (CRC).
- The method for evaluation of the impact of changes in source spatial data on the process of transformation.
- The method for evaluation of significance of changes in polygon features of spatial data.

The first method improved the method proposed by Li and Fei (2008). The author suggests using CRC calculation in order to identify changes in a feature. The CRC is an error-detecting code widely used in computer networks and storage devices in order to verify that the data were saved in full volume. The identifiers based on a set of geometrical or attributive characteristics of a feature used in the transformation process have to be created (CRC identifier). The change of these characteristics will present a change in the feature important for transformation. It means the CRC identifier will change as well. Otherwise the changed characteristics will be unimportant for the transformation process. The value of one or set attributes has to be used for detecting a change of feature attributes. However, the attribute's values alone cannot be used for creating CRC identifiers, since they will not construct unique identifiers. Additionally, a

unique identifier supported in the whole source data set has to be used. The CRC identifiers for detecting a shape or location change may be used as well.

After the research the CRC identifiers for detecting geometrical changes of the features have to be constructed as:

- for point features – ‘x of feature centroid’ AND ‘y of feature centroid’;
- for line features – ‘x of feature centroid’ AND ‘y of feature centroid’ AND ‘length of feature’;
- for polygon features – ‘x of feature centroid’ AND ‘y of feature centroid’ AND ‘length of feature’ AND ‘area of feature’.

However, using a large number of different values for the identifier is complicated, as that values are very long and differ in length. Therefore MD5 coding algorithm has to be used for CRC identifiers additionally. This algorithm is widely known as algorithm that changes values to hexadecimal values.

The features from non-updated version of spatial data set have to be joined with features from updated spatial data set. The unique identifiers of features are used for joining features. It may be original unique identifiers of features supported in a spatial data set or CRC identifiers. Following the proposed method, CRC identifiers have to be created for each feature in non-updated and updated spatial data sets. Analysis of CRC identifiers of joined features of non-updated and updated spatial data set can present changes of features. When a cardinality type between CRC identifiers of features is 1:1, it means that the feature did not change; 1:0 the feature is deleted or changed; and 0:1 the feature is new or changed. The supported original ID of features will let to detect features that changed. The feature will be changed when IDs are not changed and CRC identifiers changed. The proposed method can quickly detect all the changes, but not all changes are important for the repeated harmonisation process. Therefore the method for evaluation of the impact of changes in source spatial data on the process of transformation was proposed. The detected change in a spatial object is significant if it may affect the update process of smaller scale spatial objects:

- effect on the creation and integration of a new spatial object (*Create New Object*);
- effect on the update of an attribute of a spatial object (*Update Attribute*);
- effect on the update of the configuration (geometry) of a spatial object (*Update Configuration*);
- effect on the deletion of spatial data (*Delete Object*).

The method for the identification of changes in spatial objects affecting the update of spatial data includes a set of processes:

1. Establishing the relationship between source data before and after updating. This process is proposed for the features without unique identifiers. This step includes the creation of CRC identifiers.

2. Search for changes:

- a. Analysis of related features (*sdata.r*) and selection of the features that change after updating. Features with altered attributes fall under *Group A1* and features with a changed shape belong to *Group A2*.

- b. Selection of the features (*sdata.v2*) after updating that have no relation with the features before updating. These objects fall under *Group B*.

c. Selection of the features (*sdata.v1*) before updating that have no relation with the features after updating. These features fall under *Group C*.

3. Analysis and evaluation of features from groups A1, A2, B and C:

a. Analysis of the features from groups A1, B and C and selection of features with suitable attributes. Other features are removed from the groups.

b. Analysis of the spatial objects from group A2 and selection of features with significant changes in configuration (geometry). Other objects are removed from the groups.

c. Analysis of the spatial features from groups B and C and selection of features with suitable attributes of geometry. Other features are removed from the groups.

4. Establishment of the relationship between spatial features from groups A1, A2, B, and C and renewable data (smaller scale features) (*Change.udata.v1*).

5. Analysis of spatial features from groups A1, A2, B and C that are related to renewable data. The spatial features from groups:

a. A1 and B are identifying places for the updated attributes of features.

b. A2 identifies places for the updated configuration of features.

c. C identifies places for deletion of features.

d. Analysis of features from groups A1, A2, B and C that have no relation with renewable data. The spatial objects from groups A1, A2 and B identify places for the creation of new features.

A conception of the model is represented in figure S1.

Significant changes in configuration (geometry) have to be done in the specific order (Fig. S2).

1. Union of features from *Group A2* before and after update. Selection of changes in the geometry of the features (only part of the features is changed) (*A2.changes 1*).

2. Creation of buffer zones around source features before update (*Buffer 1*). The size of the buffer depends on the specified resolution.

3. Selection of changes in the geometry of the features inside the buffer zone (*A2.changes 2*).

4. Simplification of the changes in the geometry of the features (*Simplify Object*).

5. Creation of buffer zones around updated features (*Buffer 2*). The size of the buffer depends on the specified resolution.

6. Selection of changes in the geometry of features outside the buffer (*A2.changes 3*).

The algorithm presented in figure S2 was tested using polygon features of GDR10LT features class "Plotai" before and after updating. There were analysed 3,927 features before updating and 3,955 after updating. CRC identifiers show the changes in 1,648 features. The union of features before and after updating generated 4,916 combinations different elements. The created elements with attributes of features before and after updating identified parts of features that not changed. Otherwise created elements identified changed parts of features. Depending on proposed algorithm changed part of features were simplified and selected using buffers.

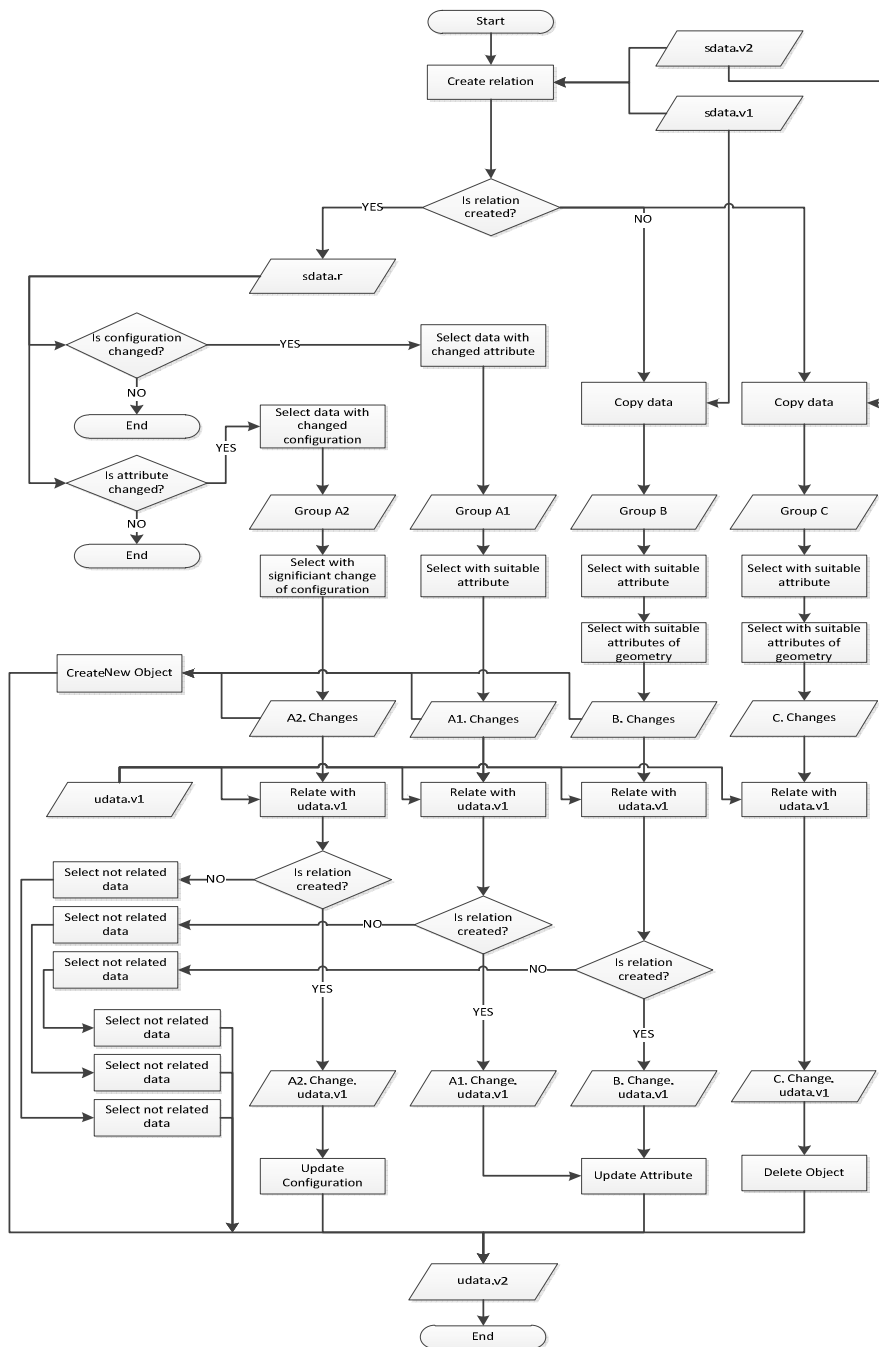


Fig. S1. Algorithm of identification of change

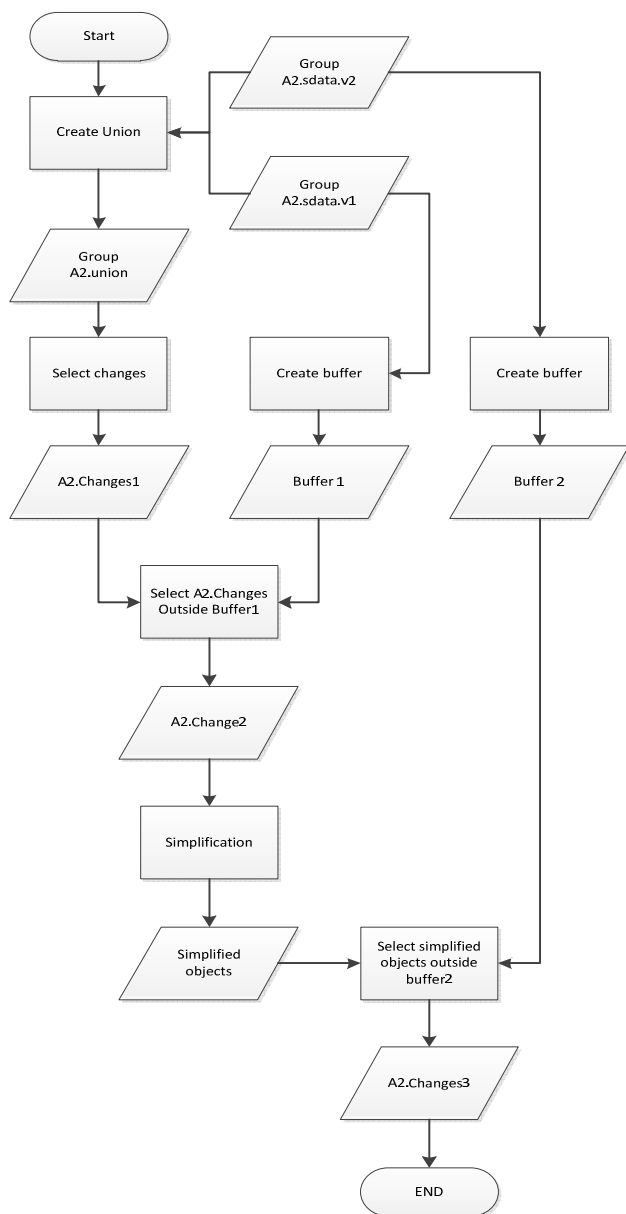


Fig. S2. Algorithm of identification of significant changes

A five-meter buffer was chosen for the identification of changes in features at the scale of 1:10,000 that influence the related features at the scale of 1:50,000. There were identified 952 changes that cross boundaries of the buffer. That indicated less than 25% of whole changes of features geometry that influence to the transformation process.

Modelling of spatial data was covered in detail in the INSPIRE Generic Conceptual Model. This document provides for detailed guidelines for the description of specifications for spatial data of INSPIRE themes. The harmonisation process allows collecting source spatial data disregarding final spatial data models and code lists used in them. Nevertheless, taking into account the description of the INSPIRE conceptual model and the methods for the identification and evaluation of changes in spatial data proposed in the thesis, it is recommended to collect the following attributes for the features when modelling spatial data, which will be used in the harmonisation process:

1. Unique identifiers of the features, which are to be supported during their life cycle.
2. Life cycle attributes of the features (date of creation, change and elimination of a feature).
3. Relations with the identical features of a final spatial data set.

The development of the unique identifiers must take into account (INSPIRE Generic... 2014) its unique character, stability, availability and suitability.

The project of the development of the Lithuanian spatial information in 2011 based on INSPIRE principles produced a Uniform (Geo)referential Spatial Data Model. This model proposes the following life cycle attributes for the features: unique identifier, beginning and end of a life cycle of a feature, version number and version date. The proposed attributes comply with the INSPIRE conceptual model (INSPIRE Generic... 2014). However, the version number is excessive, especially in case of frequent change of a feature. It is also recommended to collect, instead of a version date, the date of change of a feature as well as description and reason of the change. Such attributes would facilitate the identification and evaluation of changes in features of spatial data.

3. Development of technology for harmonisation of spatial data

The harmonisation methodology for spatial data served as the basis for introducing technology for harmonisation of spatial data of road networks, hydrography networks and land cover in GDR10LT and GDR50LT. The developed technology covers models for loading of source data to the source database, for the verification, generalisation and restructuring of the loaded data and for the verification of final spatial data. All models were built using *ArcGIS* software *ModelBuilder* possibilities. These models are able to use the technologies of the Lithuanian SDI. Different models for generalisation of individual features in classes of roads, hydrography network and land cover were developed as well. They are presented below.

Generalisation of a road network. A network of roads is represented as a set of polyline objects in GDR10LT and GDR50LT. Apart from the simplification of the features there are important to merge parallel roadways into one centre line of roads and

to eliminate short streets. The process of road network generalisation has to include these steps: (1) the elimination of the dangles; (2) creation of a new attribute field for assigning values of road hierarchy (importance) and defining the values; (3) thinning a network of the roads/streets ("Thin road network") in territories of the cities using values of road/street hierarchy; (4) splitting features at vertices ("Split lines at Vertices"). It will give better results in the next step; (5) merging parallel features into centre lines ("Merge divided roads"); (6) thinning a network of roads ("Thin road network") in other territories using values of road hierarchy; (7) collapse and simplification of small open configurations of road segments, such as traffic circles ("Collapse Road Detail"); (8) aggregation of features with the same characteristics ("Dissolve"); (9) simplification of the features ("Simplify Line"); (10) creation of a geometric network ("Create Geometric Network").

Generalisation of a hydrography network. A network of hydrography is represented as a set of polyline objects in GDR10LT and GDR50LT as well. The process of generalisation has to include these steps: (1) creation of a new attribute field for assigning values of road hierarchy (importance) and defining the values; (2) elimination of the dangles using the solution described above; (3) elimination, following the description above, of the short linear features that are adjacent to other linear features; (4) simplification and smoothing of the features ("Simplify Line", "Smooth Line"). The dangles have to be eliminated using some iteration of 2 and 3 step.

Generalisation of land cover. Land cover objects are represented as polygon features. Usually this layer represents different themes of the real world features: built-up territories, lakes, rivers, forests, agricultural areas, etc. Generalisation of land cover has to be performed following some rules: an algorithm for simplification of features and parameters depend on the type of an object; the boundaries of the neighbouring objects in the defined distance have to (or may) be identical; the boundary of the object has to (or may) be identical to the interior of a road, railway and hydrographic objects if they are in the defined distance from them. The generalisation of land cover is carried out in the following steps: (1) create boundaries from the polygon features and extend them to the centre lines of the sliver for eliminating slivers polygons; create polygons from the extended boundaries; (2) select objects with important attribute information that could not be lost during the generalisation process (e.g. name of the lake); generalise selected polygons to points ("Feature to Point"); (3) select and export objects to different layers according to its theme; (4) aggregate neighbouring objects (separately in each layer), if they are in the specified distance from each other ("Aggregate Polygons"); use option "Preserve orthogonal shape" in the case of aggregating objects, which represent anthropogenic features (e.g. built-up territories, territories of airports); (5) simplify objects separately in each layer (use "Simplify Polygons", if objects represent natural features of the real world and "Simplify Buildings" for the objects representing anthropogenic features); different generalisation parameters may be used for each layer; (6) copy attribute information from the points (the result of the 1st step) ("Spatial Join", "Calculate Field"); (7) dissolve all layers into one ("Dissolve"); (8) eliminate small objects by merging them with neighbouring objects that have the longest shared border and after that the largest area ("Eliminate"); this step must include some iterations of elimination with different values of min area of the eliminated object.

The models were tested using spatial data of GDR10LT covering 5,000 ha territory. The testing of the models was made in two stages during:

1. designing of the algorithms of generalisation for GDR10LT feature classes “Keliai”, “Hidro_I” and “Plotai” and developing of models in order to select the most effective *ArcGIS* tools, orders of processes, parameters and the number of iterations, as it is described in the methodology for spatial data harmonization.
2. testing of developed models in order to prove the harmonisation abilities of these models. The transformation and generalisation of spatial data of GDR10LT inside the test area were made at this stage.

An analysis of the generalisation results shows the reduction of features. The features representing streets changed most: the number was reduced about 38%, and a length over 16%. Mostly those were the features that represented the high density streets in urban areas. The distances between such features were smaller than in non-urban areas. Therefore the harmonisation of spatial data in different scales usually takes into account the possible minimum distances between features. During generalisation those features were thinned more than others. The generalisation of the road network eliminates short dangles of features that do not visually appear in smaller scale maps either. The features of GDR10LT representing road dangles with a length below 200 meters in countryside and 100 meters in urban areas were eliminated during generalisation. Therefore, the number of the features representing the field and forest roads decreased about 7.8%. Generalisation of land cover influences the area of the polygon feature. It may increase. It depends on the surrounding features that can be aggregated with more important neighbouring features. Simplification algorithms may influence the area of a feature as well.

The correctness of the generalised features was tested using models for evaluating final spatial data. They tested topology and attributes of the features. The results show incorrectness of generalised spatial data elements. The features smaller than available minimum were found. Such small features may be easily eliminated by aggregating them to the neighbouring features.

General conclusions

The research conducted during the dissertation led to the following general conclusions:

1. The proposed methodology for the harmonisation of spatial data allows developing the technology designed to harmonise spatial data of different scales representing land cover and a road and hydrography network. The harmonisation models for spatial data sets GDR10LT and GDR50LT were developed by using the identical software applied in the Lithuanian spatial data infrastructure.
2. An improved harmonisation methodology was prepared for spatial data providing a procedure for the identification and assessment of changes in the spatial data and specifying the assessment of a change in the shape of polygon features.

3. An improvement for spatial data modelling was proposed in order to collect information about changes in features of spatial data and their type.
4. Methods for the identification and evaluation of changes in spatial data were developed. Using these methods it is possible to identify potential areas of spatial data transformation depending on the final model of the spatial data and cartographic scale of spatial data.
5. The proposed evaluation methodology for spatial data changes additionally estimates changes in geometry of spatial data polygon features. Therefore, when transferring spatial data from a source database in the process of harmonisation, only those elements can be transferred, if necessary, which changes in features or geometry are essential to the final spatial data set, depending on the data model applied and cartographic scale. Such methods realise a partial transformation and, where necessary, generalisation of spatial data elements in the process of harmonisation of spatial information.
6. The improved algorithms for the elimination and aggregation of features, elimination of dangles and matching of neighbouring features of spatial data may be used for the generalisation of individual spatial data or their sets, as well as for the harmonisation between the different scales of spatial data in the cartographic data sets.
7. In the process of developing harmonisation technologies for different scales of spatial data sets it is necessary to set a priority sequence of spatial data features being generalised, taking into account, relationships between the features.

Priedai¹

A priedas. GDR10LT duomenų modelio ištrauka

B priedas. GDR50LT duomenų modelio ištrauka

C priedas. Modelių įgyvendinimo pavyzdžiai

D priedas. Bendraautorių sutikimai teikti publikacijose skelbtą medžiagą daktaro disertacijoje

E priedas. Autorės mokslinių publikacijų disertacijos tema kopijos

¹ Priedai pateikiami pridėtoje elektroninėje laikmenoje.

Lina PAPŠIENĖ

ERDVINĖS INFORMACIJOS MODELIAVIMO IR HARMONIZAVIMO TOBULINIMAS
SKAITMENINĖJE KARTOGRAFIJOJE

Daktaro disertacija

Technologijos mokslai,
Matavimų inžinerija (10T)

THE IMPROVEMENT OF MODELLING AND HARMONISATION OF SPATIAL
INFORMATION IN DIGITAL CARTOGRAPHY

Doctoral Dissertation

Technological Sciences
Measurement Engineering (10T)

2014 11 17. 15,0 sp. l. Tiražas 20 egz.
Vilniaus Gedimino technikos universiteto leidykla „Technika“,
Saulėtekio al. 11, 10223 Vilnius,
<http://leidykla.vgtu.lt>
Spausdino UAB „Ciklonas“
J. Jasinskio g. 15, 01111 Vilnius